Search Platforms:
Big Data and Sponsored Positions[*]

Maarten Janssen[†] Thomas Jungbauer[‡] Marcel Preuss[§] Cole Williams[¶]

October 17, 2023

Abstract

Search platforms that possess abundant consumer-specific information are ubiquitous in today's economy. We study a search platform's incentives to rank products on their website in response to a consumer query taking the strategic incentives of both firms (sellers) and consumers (buyers) into account. Rankings are important to help consumers direct their search efforts and, as a result, affect firms' sales. We adapt tools from the social learning literature to characterize the platform's optimal behavior if the number of firms to rank is large. If the platform's objective is to maximize revenues from selling a sponsored position, then it fully obfuscates organic slots, while increasing the informational content of the sponsored slot. The welfare effect of sponsored positions crucially varies with the platform's additional objectives. For example, if the platform's additional objective is to maximizes sales commissions, then the consumer benefits from sponsored positions.

In 2023, the worldwide market for digital advertising is projected to reach $679bn. Its biggest component—accounting for about 40% or $280bn—is *search advertising*.[1] Search platforms (such as Google, Tripadvisor or Yelp) assist consumers looking for a product or service. The consumer submits a keyword, the platform provides a list of results, and the consumer inspects these results in whatever order she prefers. In producing this list, the

[*]We are grateful to Simon Anderson, Heski Bar-Isaac, Marco Battaglini, Emiliano Calvano, Federico Echenique, William Fuchs, Bruno Jullien, Justin Johnson, Tony Ke, H. Tai Lam, Massimo Motta, Antonio Penta, Patrick Rey, Sandro Shelegia, Vasiliki Skreta, Steve Tadelis, Nikhil Vellodi, Michael Waldman, Mark Whitmeyer, Leeat Yariv, Yang Yu and seminar participants at the Workshop on Digital Platforms (Barcelona, 2022), The Economics of Platforms Workshop (Capri, 2023), SAM Workshop (Bristol, 2023), IIOC (Washington D.C., 2023), Tinbergen Institute Conference (Rotterdam, 2023), SAET (Paris, 2023), Strategy and Economics of Digital Markets Conference (Ithaca, 2023), EARIE (Rome, 2023) and AMES (Beijing, 2023) for their helpful comments and valuable discussions. All remaining errors are our own.

[†]Department of Economics, University of Vienna. maarten.janssen@univie.ac.at

[‡]SC Johnson Graduate School of Management, Cornell University. jungbauer@cornell.edu

[§]SC Johnson Graduate School of Management, Cornell University. preuss@cornell.edu

[¶]Department of Economics, Durham University. cole.r.williams@durham.ac.uk. Williams acknowledges financial support from the Austrian Science Fund FWF under project number FG 6-G.

[1]See https://www.statista.com/outlook/dmo/digital-advertising/worldwide. The other major digital advertising categories are audio, banner, classifieds, influencer, in-app, social media and video, none of which generates an amount of spending comparable to search advertising.

platform may draw on information about the consumer (demographics, past searches, order histories, etc.) as well as on other consumers' behavior. Search advertising refers to the practice of advertisers paying search platforms for specific positions in the list.

The power of these paid search results, or *sponsored positions*, stems from (a) the way they steer consumer search activity and (b) the platform leveraging its information about the consumer. Online search pioneers and Google founders Sergey Brin and Larry Page understood the commercial value of sponsored positions well before the advent of search advertising, warning of their implications:

> "We expect that advertising funded search engines will be inherently biased towards the advertisers and away from the need of the consumers."[2]

In this paper, we argue that the welfare effects of sponsored positions and their impact on the overall ranking of search results are significantly more complex than this quote suggests.

Building on the seminal consumer search papers of Wolinsky (1986) and Anderson and Renault (1999), we analyze the implications of search advertising for search platforms, firms (advertisers) and consumers by addressing a number of pertinent, policy relevant questions: How does the introduction of sponsored positions affect the informational content of both sponsored and organic search results? How do sponsored positions affect the social welfare generated by the platform? How does data quality shape the platform's incentives to rank items across sponsored and organic positions and firms' incentive to bid? And to what extent is the platform's ranking of alternatives affected by additional/alternative revenue streams?

In response to a representative consumer submitting a keyword, a search platform provides a ranking of a finite but large number of keyword-relevant firms. Invoking big data techniques, the platform has some understanding of the consumer's preferences over firms, represented by a *match score* for each firm that is informative of the consumer's ultimate *match value*. The platform chooses a ranking algorithm that assigns firms across slots (the sponsored and organic ones) based on their bids and match scores. In our baseline model, which we extend in several directions, firms are ex ante symmetric. They submit bids to win the sponsored slot and charge a uniform price that does not depend on the final ranking. After the ranking is presented, the consumer sequentially inspects the platform's search results in her preferred order, each time incurring a search cost to learn their match value,

[2] This quote seems to have first appeared in 1998 and later in the same year published as Brin and Page (1998) introducing the Google search engine to the wider scientific community.

until she decides to acquire a product/service of her liking or to abort search altogether.

Our main result is that if the number of firms is large, a search platform that maximizes revenue from its sponsored slot assigns the search results it deems most relevant for the consumer's keyword search to sponsored positions, while it chooses to uniformly randomly assign, i.e., *fully obfuscate*, the remaining search results across organic positions. In order to prove our main result we adapt a result from the literature on social learning to deal with the challenge of consumers learning about match scores of not yet inspected search results.

In the subsequent analysis, we show that our main result generalizes to a number of settings that incorporate important real-world features. First, we show that if the platform's revenue is based on both sales commissions and proceeds from auctioning a sponsored slot, then full obfuscation of organic positions in fact serves a dual purpose: not only does it maximize the revenue from sponsored positions but it also maximizes sales commissions. Second, we show that uniform obfuscation remains important in cases where firms do have some private information about match values. Finally, our main result continues to hold if firms charge different prices. We conclude our analysis by showing that sponsored positions may increase or decrease consumer welfare depending on the search platform's objectives.

We now explain these results in more detail. To understand why full obfuscation—i.e., minimizing information provision—of the organic slots is optimal when the number of firms is large, consider first the consumer's problem. Under full obfuscation, no matter how the consumer searches, she essentially picks a random firm once she decides to search beyond the sponsored slot. It follows that full obfuscation minimizes the consumer's utility from inspecting the organic positions, which means that the likelihood to stop searching after the sponsored slot is as high as possible. We show that the advertisers' willingness to bid for the sponsored slot is therefore also maximized, even though full obfuscation minimizes the probability that a consumer ever returns to the sponsored slot.

To see why the platform designs the algorithm so that the firm with the highest match score wins the auction for the sponsored slot, note first that in the baseline model all firms submit identical bids. As a consequence, if the platform assigns the firm with the highest match score to the sponsored slot in case of ties, the sponsored position contains informational value for both consumers and firms. Therefore, consumers optimally examine the sponsored slot first before inspecting any of the organic ones. Firms, on their part, learn that by winning the auction the platform assesses they are more likely to have the best possible match for the consumer. This increases the likelihood the consumer buys their product, thereby raising the firms' willingness to bid for the sponsored slot. A similar argument shows that if firms have private information about the consumer's match value with their product the platform leverages its match score information to maximize the match value that the consumer expects to find in the sponsored slot. It follows that it is not just a benevolent act of platforms to take match scores into account when allocating the sponsored slot: doing so makes firms want to bid more to acquire the sponsored position.

A consequence of our result is that the better a search platform is able to predict consumers' match values (in the sense of Lehmann (1988)) the more profits it makes and the larger consumer surplus. Putting the firm with the highest match score in the sponsored slot (and uniformly obfuscating the organic slots) implies that with a platform that is better informed consumers are more likely to buy from the firm in the sponsored slot. Therefore, firms will be willing to pay more for getting placed in the sponsored slot and consumers will likely get a better match and economize on their expected search cost.

A major obstacle in any analysis of search on a platform is that as the ranking of firms is strategically chosen by the platform, inspecting a search result allows consumers to make inferences about match values at firms that are not yet inspected. It is well-known that learning in consumer search models leads to complications (see, e.g., Garcia and Shelegia, 2018), and that a reservation price may fail to exist (see, e.g., Rothschild, 1974; Janssen et al., 2017). In order to overcome these difficulties we adapt the so-called *mixing property* of stochastic processes that is used in the literature on social learning (Mossel et al., 2020). We establish a Mixing Principle for Consumer Search that states that—due to the independence of match values across firms—the match value of the firm in the sponsored slot can be strongly correlated with the match values of at most a few firms in organic positions. As such, when the number of firms increases, the additional information about the continuation value of search conveyed by a low (or high) realized match value vanishes. This is important, for example, for our obfuscation result outlined above. In general, minimizing the value of continuing to search beyond the sponsored slot may not be sufficient to maximize the probability of a sale at the sponsored position due to return demand: with consumer learning, consumers do not necessarily search all organic slots once they search beyond the sponsored slot. Applying the Mixing Principle imposes bounds on how large return demand can be. By means of counterexamples we show that if the number of firms is small the optimal ranking depends on the details of the situation and, in particular, it may not be optimal to fully obfuscate the organic slots or to design an algorithm that allocates the firm with the best match score to the sponsored slot in case of ties. We believe that our Mixing Principle is fruitful for other search and inspection environments in which learning is important.

Fully obfuscating organic positions does not only maximize the advertisers' willingness to bid for the sponsored position, but also the number of organic slots a consumer inspects in expectation, thereby increasing the likelihood a consumer eventually buys a product/service through the platform, independent of whether the platform offers sponsored positions or not. This is important as in real-world markets search platforms do not only generate revenue from selling sponsored positions but often earn sales commission when successfully intermediating firms and consumers. Therefore, remarkably, when the number of firms is large, maximal obfuscation optimizes both revenue streams (from the sponsored position and sales commissions) at the same time.

Our main obfuscation result persists when adding other real-world features of search advertising. If both the platform and firms hold some consumer-relevant information, then independent of the distribution of information the platform can extract this information through the auction for the sponsored slot. Thus, our main intuition about an informative sponsored position and full obfuscation of organic slots persists. On the other hand, even if firms quote different prices for their products on the platform, the platform continues to fully obfuscate organic positions and to sell the sponsored slot to the firm with the highest match value. However, this firm is not necessarily the one the consumer prefers in expectation anymore depending on its price.

To address the welfare effects of introducing sponsored positions, we investigate how the platform ranks products in case it does not offer sponsored slots. We show that welfare effects strongly depend on the platform's alternative objectives. If, in the absence of sponsored positions, the platform cares about its reputation via consumers' ex post utility (so that the platform's and consumer interests are aligned), it would not want to use uniform obfuscation. The obfuscation of organic slots that comes with having a sponsored position harms consumers (and social welfare) in two ways: First, the expected match value of consumers who buy at the first position decreases (as consumers are more willing to stop searching). And second, consumers who continue to search beyond the first position expect to spend more time inspecting other products to achieve a given match value.

If, on the other hand, the platform maximizes sales commissions revenue, as is common practice, then (perhaps surprisingly) consumers are better off with sponsored positions. Independent of whether the platform offers a sponsored position, it wants to fully obfuscate organic slots. When there is a sponsored position, however, the platform allocates it to the firm with the best match score, thus providing the consumer with more information. Evaluating these welfare results, we conclude that sponsored positions on platforms charging commission fees (like Yelp and Booking.com) in fact help consumers find a good match more quickly. On search platforms that do not charge commission fees such as Google, by contrast, the reverse may be true.

Starting with Athey and Ellison (2011), Chen and He (2011) and Eliaz and Spiegler (2011) a growing literature on position auctions explicitly takes into account that the value of a position depends on consumers' search patterns. A key difference between our analysis and this literature is that in our paper the platform has information about consumer preferences. Thus, our paper focuses on the important policy question that has occupied policy makers and regulators of how online search platforms use their information on consumer search and purchases to steer consumer search.

More recent papers have addressed questions related to search platforms that are different from the questions we address. For example, using insights from the ordered search literature Anderson and Renault (2021) find that different rankings maximize total industry profits, total welfare or consumer surplus. Bar-Isaac and Shelegia (2022) mainly focus on how the way a platform sells a single slot (either through auction, or by means of an algorithm) affects prices and consumer welfare. Janssen and Williams (2022) study how a social influencer, whose preferences are correlated with her followers, affects market outcomes by recommending one product.[3] Motta and Penta (2022) consider how competitors bidding on brand keywords affect the competition between brands in a platform that has only one sponsored, and one organic position. Ke et al. (2022) also analyze how a platforms uses its information about consumer preferences to its advantage. In their paper there is, however, only one firm with which a consumer has a match and the platform must determine how prominently to display that firm. None of these papers considers the main questions

[3]Recommendation systems have also been studied by Che and Hörner (2018) and Glazer et al. (2021) among others. These papers, however, focus on how a recommender may act strategically in order to learn from agents' subsequent choices and study settings with identical goods. Armstrong and Zhou (2022) consider how information provision affects consumers in the absence of consumer search.

we address in this paper, which is how introducing a sponsored position affects the search platform's ranking and the effect this has on the welfare generated by the platform.

In the context of a firm directly selling to consumers Nocke and Rey (2023) find that "garbling" of information may be optimal for the seller as it induces a buyer to inspect a larger number of items before terminating search, which is related to the second role of obfuscation described above when the platform also gets revenues through sales commissions. Similar to Chen and He (2011) and Anderson and Renault (2021), however, they make the simplifying assumption that a firm-consumer pair either has a match or no match and that the consumer has the same value at any firm where there is a match. As a consequence, consumer learning about match values with firms that have not yet been inspected is not an issue in their paper (because as soon as there is a match, the consumer buys). Another important difference is that we study the interaction between sponsored and organic positions.

The effects of rankings on choices have also been studied empirically (see, e.g., Ghose et al. (2014), Ursu (2018) and Donnelly et al. (2022)). These papers show that personalized rankings affect consumer choices and induce important positive welfare effects. They only consider organic slots though and neither study the effects of sponsored positions nor the incentives of search platforms to provide alternative rankings.

The remainder of this paper is organized as follows. Section 1 below introduces the main model in which the platform maximizes its revenue from selling the sponsored slot. Section 2 then presents our main result, the intuition of the underlying proof and discusses the quality of platform information. Section 3 generalizes our main result to scenarios where (i) the platform earns revenue from sales commissions, (ii) both the platform and the firms hold consumer-relevant information and (iii) firms have heterogeneous prices. In Section 4 we build on our formal results and discuss the welfare effects of sponsored positions and how they depend on the platform's objective. Finally, Section 5 concludes with a discussion.

1 The Model

The market comprises a platform, n firms selling a horizontally differentiated product, and a representative consumer. The consumer demands one unit of the product and has an unobserved match value v_i with firm i. Match values are independently and identically distributed across firms according to a continuous distribution G with (a possibly unbounded) support $[\underline{v}, \bar{v}] \subset \mathbb{R}$. The platform's information regarding the consumer's match value with firm i is summarized by a score θ_i, following a continuous distribution $F(\theta_i|v_i)$ with compact support $[\underline{\theta}, \bar{\theta}] \subset \mathbb{R}$. Higher scores indicate "good news" in the sense that $F(\cdot|v_i')$ has likelihood ratio dominance over $F(\cdot|v_i)$ if $v_i' \geq v_i$. Let z denote an independent non-atomic random variable with support Z that the platform can use as a randomization device to play a mixed strategy. We denote the probability measure on $\Omega = [\underline{\theta}, \bar{\theta}]^n \times [\underline{v}, \bar{v}]^n \times Z$ by μ and denote probabilities and expectations with respect to μ by $\mathbb{P}[\cdot]$ and $\mathbb{E}[\cdot]$, respectively.

The platform displays a *ranking* $x \in X$ of the firms to the consumer, where X denotes the set of firm permutations.[4] The platform forms the ranking by running an auction in which firms submit bids, the winner of the auction is placed at the top in the "sponsored" position (i.e. $x(i) = 1$ implies i is sponsored), and all other firms are arranged into the remaining "organic" positions. Denote firm i's bid by $b_i \geq 0$. Given the bids and scores, the platform's algorithm determines which firm wins the sponsored position and how the remaining firms are arranged into the organic positions. The platform's *algorithm* is a function mapping bids, scores, and realizations of the randomization device to rankings, $a : \mathbb{R}_+^n \times [\underline{\theta}, \bar{\theta}]^n \times Z \to X$. Let $\mathcal{A}_n$ denote the set of (measurable) algorithms in the game with n firms. Denoting the vector of bids by $\mathbf{b} = (b_1, b_2, \ldots, b_n)$, the platform also specifies a *payment rule* $\rho : \mathbb{R}_+^n \to \mathbb{R}_+^n$ with $\rho(\mathbf{b}) = (\rho_1(\mathbf{b}), \ldots, \rho_n(\mathbf{b}))$ whereby $\rho_i(\mathbf{b}) \leq b_i$ is the amount Firm i pays the platform conditional on it winning the sponsored position. Let $\mathcal{P}_n$ denote the space of (measurable) payment rules in the game with n firms.

Notice that the algorithm and payment rule jointly define the structure of the sponsored search auction. For example, an algorithm that always places a firm with the highest bid in the sponsored position and a payment rule satisfying $\rho_i(\mathbf{b}) = b_i$ correspond to a standard first-price auction. More generally, the platform can use the information contained in the scores to determine the winner of the auction.

Consumers are initially uninformed of their match values with firms and can only find this out through costly sequential search. At each point along the search path, the consumer can select any position in the ranking and incur the inspection cost $s > 0$ to learn the match value with the firm located in that position, buy the good from a firm whose match value the consumer has already inspected, or exit the market and take an outside option of zero. When searching they only observe their match values with different firms and for example not the platform's match score. In principle, the consumer search problem is like Pandora's box

[4] That is, X is the set of bijections from $\{1, \ldots, n\}$ to itself.

problem as studied in Weitzman (1979). However, knowledge that the platform utilizes in a ranking algorithm creates interdependence between match values so that inspecting the goods of one firm provides the consumer with information about other firms. Consumers have perfect recall when searching.

The timing of the interaction is as follows. First, the platform commits to an algorithm and payment rule which is observed by the firms and consumer. Second, firms privately submit their bids to the platform. Third, Nature determines the match scores and values. Fourth, the platform receives the firms' scores and bids and the algorithm determines the position each firm takes in the list. The consumer receives the list and then proceeds with her search. At each point along the search path, the consumer's information consists of the the algorithm, payment rule, as well as the realized match values at all positions she has previously inspected.

The consumer's payoff is equal to the match value minus the price of a good she purchases net the search costs. A firm's profit equals the revenue minus product cost and any fee paid for the sponsored position. Unless explicitly stated otherwise (as in Subsections 3.3 and 4.3) we assume that all firms charge price p and normalize their production costs to 0. The platform's expected profit corresponds to the expected revenue from the sponsored search auction, but in later sections we also consider alternative objectives. We focus on symmetric Perfect Bayesian Equilibria.

We end this section with a few comments on the model. First, we assume that the platform commits to its ranking algorithm. We see this as a reasonable approximation of the real world situation where platforms submit a ranking of alternatives within a split second after the consumer has typed its key words. Consumers typically use the same platform over and over and see the resulting rankings and how they satisfy their needs. Platforms may, of course, work on different algorithms to improve their functioning, but will implement new algorithms only once in a while. Without commitment, other outcomes than the ones we focus on in this paper may be supported. For example, if firms and consumers believe that the platform's ranking, including the sponsored position, is completely random, then the platform may not be able to do better than indeed randomly allocating firms to positions.

Second, the questions we address are akin to the ones studied in the literature on information design, as initiated by Kamenica and Gentzkow (2011), in the sense that the platform chooses which information to release to firms and consumers. However, unlike most of that literature, the platform's choice set is limited in that it can only choose a ranking of alternatives and it should give the same information to both consumers and firms.

Third, we implicitly assume that the platform has to rank all firms and cannot present a truncated ranking. We see this as a short-hand for a platform not having unlimited market power. In particular, platforms may prefer to rank all firms and not run the risk of consumers not using their services in the future, if consumers would not continue to use the platform if a firm they expect on the list is not ranked at all. The platform may also directly benefit from organic search results, for example, from commission fees they generate; see Section 3.1.

Fourth, the model treats firms' prices as exogenous. In particular, prices do not depend on whether or not a firm is recommended. We think that this is realistic in many cases in which the revenue a firm makes is only to a limited extent dependent on the sales via the search platform. Implicitly, we also assume that all firms charge identical prices, but that turns out to be inessential as we will explain in the next section after stating our main result.

2 Obfuscation and the Revenue from Sponsored Positions

In this section we state and explain our main result: as the number of firms grows large, it is optimal for the platform to fully obfuscate the organic slots so that they dot not contain any information regarding the platform's match scores, and to allocate the sponsored position to the firm with the highest match score. By means of examples, we show that for small n the optimal algorithm depends on details of the environment and that it may *not* be optimal for the platform to obfuscate organic slots or to put the firm with the highest match score in the sponsored position.

We first introduce two definitions. Let $\Pi(a)$ denote the platform's expected profit when the consumer and firms play an equilibrium of the subgame following the selection of a. In general, an algorithm could induce multiple equilibria that differ in the amounts firms bid for the sponsored slots and in that case $\Pi(a)$ simply selects the pay-off of an arbitrary equilibrium in that set. In what follows, we focus on a sequence of algorithms $\{a_n\}_{n\in\mathbb{N}}$ where $a_n \in \mathcal{A}_n$ for all $n \in \mathbb{N}$.

Definition 1. *A sequence of algorithms $\{a_n\}_{n\in\mathbb{N}}$ is* ***asymptotically optimal*** *if for every sequence $\{a'_n\}_{n\in\mathbb{N}}$ and $\epsilon > 0$ there exists an n^* such that $n \geq n^*$ implies $\Pi(a_n) + \epsilon > \Pi(a'_n)$.*

Definition 2. *An algorithm $a \in \mathcal{A}_n$ exhibits* ***uniform obfuscation*** *if the firms that lose the* *auction are assigned to each of the organic positions with uniform probability.*

Using these definitions, we can now state our main result.

Theorem 1. *There is a sequence of uniformly obfuscating algorithms that is asymptotically optimal. The optimal algorithm allocates the sponsored slot to the firm with the highest match score.*

Implementing a sponsored search auction and a ranking algorithm introduces interdependence between the consumer's conditional match values across firms: consumers may use observed match values to make inferences about the match values with firms they have not yet inspected. The ability to influence the consumer's learning over the course of search introduces a strategic tension in the platform's objective of designing the algorithm. On one hand, the sponsored firm's initial demand can be made larger by providing a less informative ranking of the organic positions because this reduces the consumer's expected payoff from continuing search beyond the sponsored firm. On the other hand, supplying some information in the organic ranking can bolster the sponsored firm's return demand as observing a low match value at an organic firm makes the consumer pessimistic about the remaining organic firms, making returning back to the sponsored firm more attractive. In addition, for certain realizations of match scores, the platform may prefer not to put the firm with the highest match score in the sponsored position as the consumer may become more optimistic about finding an even better match value (and thus continues to search on) if the platform would put the highest match score first. Hiding information to the consumer may then increase the probability that the consumer immediately buys from the sponsored slot.

Our main result provides a clear picture of what the optimal algorithm is by focusing on the many real-world applications where the number of potentially relevant firms for a search query is large. This allows us to adapt the the property of *mixing* (Mossel, Mueller-Frank, Sly, and Tamuz, 2020, Lemma 1) to establish our main results. In order to develop the intuition for why the result holds for large n, we start by presenting two examples showing that, due to learning effects, our main result may fail to hold with a small number of firms. The first example shows that uniform obfuscation may fail to be optimal, while the second example shows that the platform may not want to put the firm with the best match score in the sponsored position. The second example also illustrates that with a small number of firms if the platform always puts the firm with the highest match score in the sponsored position, consumers may have non-monotonic reservation values: they continue to search for intermediate match values, but stop searching for high or low match values.

Example 1. Suppose there are three firms. The consumer's match value is either low ℓ, medium m, or high h and a product is only worth purchasing if it provides at least a medium value. A firm's match score is L when the value is low and H when the value is either medium or high, i.e., the platform can distinguish firms with low match scores from other firms, but cannot distinguish firms with medium and high match scores.[5] Suppose the platform employs the following algorithm. The firm with the highest bid is placed in the sponsored position, ties are broken in favor of the firm with the highest match score and further ties are broken with equal probability. For the two nonsponsored firms, if only one of them has a high signal it is placed in the second position with probability $\alpha \geq \frac{1}{2}$, otherwise they are arranged in the organic positions with equal probability. Uniform obfuscation is a special case where $\alpha = \frac{1}{2}$.

Given the algorithm, the consumer's optimal search proceeds in the following manner. If the sponsored firm's value is high h, then the consumer buys it immediately since there is no advantage from continuing. If instead the sponsored firm's value is low ℓ, then given the algorithm, the consumer learns that all remaining firms must likewise have low match values and so the consumer might as well exit the market. If, however, the consumer observes m in the sponsored slot, then it might still be prudent to continue searching as some remaining firm might deliver a higher match value. In Appendix S.2, we show that one can find parameter values such that (*i*) if $\alpha = \frac{1}{2}$, the consumer continues searching when observing m in the first and the second search, but halts otherwise, and that (*ii*) for some values of $\alpha > \frac{1}{2}$ the consumer will inspect the second slot if the sponsored position provides a medium match value, but will not find it optimal to search further. Note that by providing some information in the organic slots, the nonuniformly obfuscating algorithm makes inspecting the second firm more desirable, but increases a sponsored firm's return demand as consumers will not inspect the third firm. In addition, it will also affect the expected profits of a non-sponsored slot. Appendix S.2 shows that firms are willing to bid more to get the sponsored slot instead of an organic slot with an algorithm that is not uniformly obfuscating the organic slots, i.e.,

[5]This example departs from the assumptions of our model in that the distribution of match values conditional on the match scores do not share the same support. This is insignificant to the particular example since we could modify the distributions to $\mathbb{P}(\{\ell\}|L) = \mathbb{P}(\{m, h\}|H) = 1 - \varepsilon$ so that the conclusion continues to hold for $\varepsilon > 0$ sufficiently small.

$\alpha > \frac{1}{2}$ as the effect on return demand dominates the other effects.

Example 2. Consider a similar example as before where there are three distinct match values l, m, h and where the platform cannot distinguish medium and high values, but perfectly recognizes a low value firm. Let there be four firms, however, and every product is, in principle, worth buying (that is, better than the outside option). Consider two different algorithms, one where the platform always puts the firm with the highest match score in the sponsored position and uniformly obfuscates the organic positions and another, second, algorithm where the platform almost always follows the same algorithm apart from the case where two match scores are L and two are H. In that case the platform puts a firm with the lowest match score in the sponsored position and uniformly obfuscates the organic positions.

In Appendix S.3 we show that there are parameters under which the second algorithm gives the winning firm a higher probability of selling. In particular, with the first algorithm the consumer buys immediately from the sponsored position whenever it contains an h or l value, but he continues to search if he sees an m value. On the other hand, with the second algorithm the consumer buys immediately from the sponsored position whatever its match value. As the second algorithm makes sure that the consumer will always buy from the sponsored position and never buys from an organic position, it is clear that the platform gets its highest possible profit as firms bid maximally to get into the sponsored position.

The main idea exploited in Appendix S.3 is that under the first algorithm when consumers find an l in the sponsored slot, they know that all remaining slots must contain l and therefore they buy immediately. The second algorithm exploits this pessimism: even though the consumer knows that after observing an l in the sponsored slot the realized match scores may now be either $\{L, L, L, L\}$ or $\{L, L, H, H\}$, the consumer may still buy immediately if the ex ante probability of L is sufficiently high.[6] The effect of this is that when the consumer observes an m value on the first search, under the second algorithm he is more pessimistic about finding an h value when continuing to search than under the first algorithm. Thus, he may stop searching under the second algorithm, but not under the first.

Proof Outline. We are now ready to convey the main elements of the proof of Theorem 1. Define the function $d_i : \mathbb{R} \times \Delta\Omega \to \mathbb{R}$ by $d_i(v, \lambda) = s - \int_\Omega \max\{v_i - v, 0\} d\lambda(\omega)$, which

[6] To some extent, this is reminiscent of Bayesian Persuasion as the platform pools good and bad events subject to the constraint that the consumer stops searching immediately.

reflects the difference between selecting a product with value v and paying the search cost s to inspect firm i and then taking the larger of the two values when the distribution over the state is λ. Observe that $d_i(v, \lambda)$ is strictly increasing in v whenever $\max \operatorname{supp} \lambda > v$ and takes the value of s otherwise. The *ex ante reservation price* is the unique value $\bar{r}$ satisfying $d_i(\bar{r}, \mu) = 0$. Given only the prior, the consumer is indifferent between taking a good with value $\bar{r}$ and first inspecting firm i and then taking the larger value between the two.

A key idea in the argument is that when there are many firms, then for *any* possible algorithm, when the consumer finds that his match value with the sponsored firm lies below $\bar{r}$, he almost certainly has a better option for how to proceed with his search than to buy the sponsored firm's product. The tool we use to formalize this idea is the fact that independently and identically distributed (IID) random variables have the property of *mixing* (Mossel, Mueller-Frank, Sly, and Tamuz, 2020, Lemma 1). Intuitively, mixing means that any event E defined on the same probability space of a sequence of IID random variables $\{Y_i\}_{i \in \mathbb{N}}$ can only be strongly related to a finite number of them.

Lemma 1 (Mixing Principle for Consumer Search)**.** *Suppose $\bar{v} < \infty$. Consider a collection of events $\{E_n\}_{n \in \mathbb{N}}$ in Ω such that $\mathbb{P}(E_n) > \alpha$ for some $\alpha > 0$. For every $v \in \mathbb{R}$ and $\delta > 0$, there exists an $n^* \in \mathbb{N}$ such that, if $n \geq n^*$ then $|d_i(v, \mu) - d_i(v, \mu(\cdot|E_n))| < \delta$ for some firm $i \leq n$.*

Using the Mixing Principle for Consumer Search, we derive two important implications for platform profit. First, we use it to develop an upper bound on the profit attainable from using any algorithm (Lemma B.1). Second, we use the principle to show how uniform obfuscation achieves this upper bound (Lemma B.2).

For the intuition of the first claim, we argue that the probability the consumer buys from the sponsored firm when it offers a match value less than $\bar{r}$ vanishes as the number n of firms grows large. To see why this is true, take any $\epsilon > 0$ and let E_n be the event in which the consumers buys from the sponsored firm and it delivers a match value that is less than $\bar{r} - \epsilon$ when there are n firms. If the probability of this event fails to vanish as the number of firms grows large, then eventually there will be some position with an arbitrarily high probability of containing a firm $i \leq n$ for whom $d_i(\bar{r} - \epsilon, \mu(\cdot|E_n)) < 0$. But this implies a contradiction as the consumer would, on average, rather inspect some other position at the decision node in which it buys from the sponsored firm. Thus, the limiting demand facing the sponsored firm cannot be larger than the probability that its match value exceeds $\bar{r}$, which is bounded above by $1 - G(\bar{r}|\bar{\theta})$. Consequently, because the platform's profit cannot exceed the amount accrued by the sponsored firm, an upper bound on platform profit is $p(1 - G(\bar{r}|\bar{\theta}))$.

For the second claim, we next argue that when n grows large, an algorithm for which the firm with the highest match score wins the sponsored position and uniformly obfuscates the organic slots and that allocates the sponsored slot to the firm with the highest match score gives the firm in the sponsored slot a demand that is arbitrary close to this upper bound. The algorithm has two features that are important. First, by allocating the sponsored slot to the firm with the highest match score, the platform makes it attractive for consumers to start their search at the sponsored slot. When n grows large, there is almost surely a firm that has a match score that is arbitrarily close to $\bar{\theta}$. Second, even for large n, the platform could in principle choose a ranking that allows consumers to learn, but by uniformly obfuscating the organic slots, it effectuates that consumers do not learn anything from observing their match value at the sponsored slot and they expect a randomly selected firm to have a match value close to $\bar{r}$. This makes it least attractive for consumers to continue searching beyond the sponsored slot and from the above we know that the return demand is arbitrarily small for large n. Thus, it is *not* the case that when n grows large consumers cannot learn from the platform's ranking, but the platform cannot gain from rankings that do allow consumers to infer information about match values at firms that are not inspected yet. Together, the above two features ensure that a firm's demand in the sponsored slot approaches $1 - G(\bar{r}|\bar{\theta})$.

The last step of the proof argues that for large n there exists an equilibrium where firms find it optimal to bid an amount in the bidding stage that is close to $p(1 - G(\bar{r}|\bar{\theta}))$ realizing an expected profit for the platform close to the maximal attainable profits. This step has two parts. We intuitively discuss here the argument where firms do not have private information on which to condition their bid so that in a symmetric equilibrium the platform receives identical bids from all firms. Section 3.2 shows that the result actually holds for an arbitrary distribution of private information about the consumer between firms and the platform. So, suppose the platform's *algorithm* $a : B \times \Theta \times Z \to X$ allocates the sponsored position to the firm with the highest bid and in case of a tie allocates the position to the firm with the highest match scores among the firms with the highest bid. Besides, let the payment rule $\rho : B \to \mathbb{R}_+^n$ be equivalent to that of a second-price auction. Then, in an equilibrium firms do not want to outbid their competitors, because if they do they win the auction because of their higher bid and not because of their higher match score. Outbidding results in the informational value of winning the auction being lost and a firm's expected demand in the sponsored slot dropping $1 - G(\bar{r})$ (Lemma B.3). The second part characterizes the set of possible equilibrium bids. To this end, suppose all firms bid β_n. A firm's expected profit from likewise bidding β_n is equal to

$$\frac{1}{n}(\pi(1, a_n) - \beta_n) + \sum_{m=2}^{n} \frac{1}{n}\pi(m, a_n),$$

where $\pi(m, a_n)$ is the firm's expected profit from playing the tentative equilibrium strategy and being relegated to the mth position. Deviating to a lower bid yields an expected profit of

$$\sum_{m=2}^{n} \frac{1}{n-1}\tilde{\pi}(m, a_n),$$

where $\tilde{\pi}(m, a_n)$ is a firm's expected profit from being in a position $m \geq 2$ following the deviation. Deviating to a higher bid yields an expected profit of

$$\hat{\pi}(1, a_n) - \beta_n,$$

where $\hat{\pi}(1, a_n)$ denotes the firm's expected profit in the sponsored position following the upward deviation. Combining these expressions, we find bidding β_n to be a best reply, if and only if,

$$\frac{n}{n-1}\hat{\pi}(1, a_n) - \frac{1}{n-1}\sum_{m=1}^{n} \pi(m, a_n) \leq \beta_n \leq \sum_{m=1}^{n} \pi(m, a_n) - \frac{n}{n-1}\sum_{m=2}^{n} \tilde{\pi}(m, a_n).$$

The proof of the Theorem shows that for large n the LHS of this expression approaches $p(1-G(\bar{r}))$, while the RHS approaches $p(1-G(\bar{r}|\bar{\theta}))$. Thus, for large n there is a continuum of equilibrium bids. By choosing a reserve price that is close to the highest equilibrium bid, the platform can easily resolve this equilibrium multiplicity to its own advantage achieving a profit close to $p(1 - G(\bar{r}|\bar{\theta}))$.

Theorem 1 abstracts from the fact that a consumer typically has the ability to leave the platform and continue searching elsewhere if they become dissatisfied. Interestingly, as long as the outside option does not preclude the consumer from visiting the platform for every algorithm, as the algorithm detailed in Theorem 1 places the best firm in the sponsored position, it continues to lure the consumers to first examine the sponsored firm even if they then continue their search elsewhere. Formally, we refer to the outside option $\eta \in \mathbb{R}$ as the payoff received by the consumer if they decide to exit the platform at any point. We then have the following result.[7]

Corollary 1. *The sequence of uniformly obfuscating algorithms identified in Theorem 1 remains asymptotically optimal for every value of the outside option* $\eta \in \mathbb{R}$.

2.1 Improved Platform Information

Now that we better understand for which purposes the platform may (not) use the information it possesses, we can also answer the question how the market is affected by the platform having more accurate information. Consider an improvement in the quality of the platform's information in the sense of Lehmann (1988). That is, if $F(\theta_i|v_i)$ is the platform's initial score distribution, then the platform has better information if its new score distribution $\tilde{F}(\theta_i|v_i)$ is such that

$$\tilde{F}^{-1}\left(F(\theta_i|v_i)|v_i\right)$$

is nondecreasing in v_i for all θ_i.[8]

The next proposition argues that both the platform and consumers are better off if the platform has better information.

Proposition 1. *A sequence of uniformly obfuscating algorithms that place the firm with the highest match score in the sponsored position leads to higher limiting profits and more consumer surplus, the better is the quality of the platform's information.*

Consequently, even though the platform's ranking is uninformative beyond the sponsored slot, it has an incentive to acquire better and better information. The more accurate the information about match values that the platform uses to decide which firm is placed in the sponsored slot, the higher is the probability that consumers buy from the sponsored slot and, thus, the higher is the firms' willingness to pay for this event.

Proposition 1 is consistent with the interpretation of search platforms being a critical gatekeeper in online markets. Basically, the platform sells "preferred access" to consumers to firms. This access is, of course, more valuable to a firm if the likelihood that the consumer likes the firm's product is higher. This is why more accurate information increases the platform's profits.

[7] The proof of the Corollary 1 is omitted as it follows from exactly the same arguments as the arguments for Theorem 1 where we replace the upper bound on sales by the sponsored firm with $1 - G(\max\{\bar{r}, \eta\}|\bar{\theta})$.

[8] Dewatripont et al. (1999) and Persico (2000) discuss economic applications of Lehmann information.

Interestingly, consumers are also better off as they are more likely to find a product they will like enough so that they will not continue searching. They benefit from improved information both through a higher expected match value conditional on buying and through a reduced expected total search cost. Firms may be slightly worse off as the ones with an organic slots are less likely to sell (even though for large n, their profits were anyway already small to begin with) and the firm in the sponsored slot has paid so much more that he is indifferent between winning and losing.

3 Generalizations

In this Section we show that our main result continues to hold when we take into account different real-world considerations. In particular, we show that our result is robust to (i) the platform's revenues also comprising sales commissions, (ii) firms also having some information regarding consumers' match values, or (iii) firms charging different prices.

3.1 Sales Commissions

In this subsection we amend the base model and analyze how the platform allocates firms across slots if it earns commission fees, i.e., a fee firms pay in the case of a transaction. Commission fees are only one of several reasons why a platform may want to rank also all firms that do not win the sponsored slot.[9] To see what commission fees add to the analysis, consider first that the platform maximizes sales commission only and that it does not have sponsored positions. In the base model with equal prices, this boils down to the platform maximizing the probability of a sale of any one product.

Note, that under the absence of a sponsored slot, the consumer does not learn about the match value of the firms in the remaining slots if the platform commits to full obfuscation. This directly follows from the independence of match values across firms. Thus, the posterior match value distribution at any stage of the search process is given by G. With an outside option of $\eta \in \mathbb{R}$, it follows that under uniform obfuscation, the consumer opts for the outside

[9]Other potential reasons include competitive pressure or advertising revenue the platform may accrue if consumers spend more time on the platform. On a broader level, the analysis of commission fees may be viewed as a proxy for other reasons why the platform cares about the organic slots as well.

option only if she inspected every slot provided that

$$\int_{\eta}^{\infty}(v-\eta)\mathrm{d}G(v) > s, \tag{1}$$

because the option value of sampling any slot is positive if this inequality holds. Conversely, if inequality (1) fails, the consumer does not even begin to search, implying a sales probability of zero if the platform uniformly obfuscates.

Recall that the consumer's reservation value $\bar{r}$ if the platform uniformly obfuscates satisfies

$$\int_{\bar{r}}^{\infty}(v-\eta)\mathrm{d}G(v) = s. \tag{2}$$

Thus, inequality (1) is equivalent to $\bar{r} \geq \eta$ so that uniform obfuscation maximizes sales commission revenue for any n, if and only if, $\bar{r} \geq \eta$.[10] As far as the organic slots are concerned the optimality of uniform obfuscation is in line with Theorem 1. Now consider that the platform derives revenue from both a sponsored slot and from commission fees. The above result on uniform obfuscation if the platform only cares about sales commissions cannot be directly applied if the platform maximizes revenue from both sources. This is because independence across slots fails to persist if the firm with the highest match score is allocated the sponsored slot. Nevertheless, as the mixing property implies that an event can only be strongly related to a finite number of IID random variables, it is true that, if $\eta < \bar{r}$, consumers should not abandon their search without a purchase as long as there is still a large number of slots remaining. Thus, the following result holds:

Proposition 2. *Suppose the platform sells the sponsored slot and earns a commission fee when selling a product. Then, if $\eta < \bar{r}$, there is a sequence of uniformly obfuscating algorithms that place the firm with the highest match score in the sponsored position that is asymptotically optimal.*

Thus, uniform obfuscation serves the dual role of maximizing revenue from the sponsored slot (by increasing the probability consumers buy from that slot and thereby increasing the bids of firms) as well as the revenue from sales commissions (by increasing

[10] It is worth highlighting that the probability of a sale is not necessarily zero if $\bar{r} < \eta$. Instead of uniformly obfuscating over slots, the platform could, for example, use a perfect ranking and in that case the value of opening some slots would potentially exceed η.

the probability the consumer buys from an organic slot, given it does not buy from the sponsored slot). This second, but not the first, role of obfuscation is reminiscent of results in Nocke and Rey (2023) who show that a multi-product firm may want to use obfuscation of product orderings to maximize profits. Note, however, that the settings are very different in that, in their setting, consumer learning is severely restricted as all products have the same value so that a consumer never continues searching if a match occurs.

If, on the other hand, $\eta > \bar{r}$, then consumers will not search if the platform uniformly obfuscates the firms over all slots. Instead, it may select a group of firms with higher match scores, indicate this is a "*premium set*" of firms, so that consumers can choose to only search among firms belonging to this set.[11] The platform may then find it optimal to uniformly obfuscate within the premium set and choose the premium set as large as possible such that consumers just prefer to search among these firms to the outside option. The gist of Proposition 2 remains: the platform puts the firm with the highest match score in the sponsored position and uniformly obfuscates other positions, but now only over a restricted set of firms. Note that the platform will typically find it optimal to include firms in the premium set that individually do not have a match score that is high enough for a consumer to want to inspect that firm. By obfuscating within a premium set, the platform can enlarge this set so as to increase the probability of a sale.

3.2 Privately Informed Firms

We have so far assumed that only the platform has access to relevant information regarding consumer preferences, and firms do not possess such information. We think this is relevant in many instances where firms do not have the relevant technology in place to digest large amounts of information. However, there are other instances where firms also do have relevant information in addition to the platform, perhaps from other sales channels. In this subsection, we therefore consider the situation where firms also have some information about how well their product fits a particular search query.

To model those instances, suppose that in addition to the platform receiving a match score $\theta_i \in [\underline{\theta}, \bar{\theta}]$, each firm receives a private signal $t_i \in [\underline{t}, \bar{t}]$ and that the consumer's match value with firm i is drawn from a distribution $G(v_i|t_i + \theta_i)$. The match scores t_i are independently and identically distributed (IID) across firms according to a compactly

[11] One way to create a premium set is to indicate some firms with a star or by putting them on a first page (with endogenous length of search results as is common on Amazon).

supported, atomless distribution $F(t_i)$. Denoting $y_i = t_i + \theta_i$, we assume, similar to the main model, that $G(v_i|y_i')$ has likelihood ratio dominance over $G(v_i|y_i)$ whenever $y_i' \geq y_i$. For technical convenience, we take $\bar{v}$ to be finite.

A firm's *adjusted bid* $\psi(b_i, \theta_i)$ is a smooth and strictly increasing function of firm's bid on $[0, p]$ and score on $[\underline{\theta}, \bar{\theta}]$. Consider an algorithm which awards the sponsored slot to the firm with the highest adjusted bid and uniformly and randomly assigns all other firms to organic positions. Normalize the smallest adjusted bid to $\psi(0, \underline{\theta}) = 0$. We suppose that for all bids $b_i > p$ the adjusted score is the same as if they had bid zero $\psi(b_i, \theta_i) = \psi(0, \theta_i)$. Specify the payment rule, ρ_n, so that the firm that wins the auction pays the value of its bid.[12] As bidding higher than p is dominated by bidding zero, it is without loss of generality to restrict the firms' strategy space to bids in $[0, p]$. Because the distribution of θ_i is atomless, the probability of a tie is zero. Thus, unlike the main model, each firm's expected profit is continuous in the bids on $[0, p]$.

To state the main proposition of this subsection let σ^n be an equilibrium strategy profile for the game featuring $n \geq 2$ firms. As the equilibrium of the game following the platform's choice of algorithm may be in mixed strategies, let B_n denote the random variable that corresponds to the winning bid in the equilibrium σ^n. Denote $\bar{b} \equiv p(1 - G(\bar{r}|\bar{t} + \bar{\theta}))$ as the expected profit a firm with private signal $\bar{t}$ and a platform's signal $\bar{\theta}$ would make if the consumer would visit that firm at his first search and decide to engage in optimal sequential search afterwards.

We can then state the following proposition, which is the analogue of our main result for the case where firms also have some private information.

Proposition 3. *There is a sequence of uniformly obfuscating algorithms that is asymptotically optimal.*

It is not difficult to see that if n is large and the combined information of the winning firm and the platform about the match score is the same as the platform's information in the main model, the platform earns the same profits, whether or not the firms have private information. The reason is that if firms do not have information as in the main model, they

[12]Note that this marks a minor difference with the main result where we considered a second-price auction. Second-price auctions are a bit more difficult to handle as when firms are ranked according to their adjusted bid, the second-highest adjusted bid may have a bid that is higher than the highest-ranked adjusted bid. When n is large, this difference becomes negligible, however, as the second-highest (adjusted) bid is arbitrarily close to the highest (adjusted) bid.

know that the platform uses match scores to allocate the sponsored slots and take this already into account when making the bid. If, on the other hand, firms have some private information in the form of a match score they realize that the firm with the highest combined match score $\bar{t} + \bar{\theta}$ will win the auction and that as far as the allocation of organic slots is concerned the platform ignores all information. Thus, for a given total match score of the winner, sales will be independent of whether firms have private information.

3.3 Heterogeneous Prices

The main model not only assumes that prices are exogenous, but also that they are the same across firms. For an analysis of the welfare consequences of sponsored positions it is important to acknowledge, however, that different firms are likely to have different prices and in this section, we generalize our main result and accommodate this fact. To keep the analysis tractable, we assume there is a finite, but rich set of prices charged by different firms. Moreover, and in line with considering that the possible sale through the search query is relatively small relative to the overall sales of a firm, we continue assuming that prices are determined independently of the outcome of the firm's ranking in this search query.[13]

Together with allowing for different firms charging different prices, we also introduce different cost levels. Thus, we introduce a finite set of production costs $C = \{c^1, c^2, ..., c^L\}$ and a finite set of prices $\mathcal{P} = \{p^1, p^2, ..., p^K\}$. We denote the joint distribution over $C \times \mathcal{P}$ by $H(\cdot,\cdot)$ with probability mass function $h(\cdot,\cdot)$. For this setting, we have to re-define the ex ante reservation value $\bar{r}$ as the unique solution to

$$\sum_{p\in\mathcal{P}} h_p(p) \left(\int_{\bar{r}+p}^{\bar{v}} (v - p - \bar{r})\, \mathrm{d}G(v) \right) = s, \tag{3}$$

where $h_p(\cdot)$ denotes the unconditional probability mass function over prices in $\mathcal{P}$. Thus, we can write the ex ante expected profit of a firm in the sponsored slot with cost c_j, price p_j and

[13] Janssen and Williams (2022) consider a model where a social influencer recommends followers to consider a certain product. In this context, they consider that firms may change their prices depending on whether or not they are recommended. Their analysis suggests that the conclusions we derive here could be extended to situations where firms' prices are endogenous. In particular, a firm that wins a sponsored slot (as the recommended firm in Janssen and Williams (2022)) will optimally adjust its price (upwards) in response to the favorable news of being awarded the sponsored slot. This will further increase the willingness to bid to get it and boost the platform's profits.

match score θ when the platform uniformly obfuscates over all organic slots and $n \to \infty$ as

$$(p_j - c_j)\left(1 - G(\bar{r} + p_j|\theta)\right). \tag{4}$$

Let $\left(c(\theta), p(\theta)\right)$ be the cost-price pair that maximizes this profit for a given match score $\bar{\theta}$ subject to consumers to prefer clicking on the sponsored slot over clicking on any organic slot under uniform obfuscation. For simplicity, throughout this section we assume that the set of combinations of production costs and prices, $\mathcal{C} \times \mathcal{P}$, is sufficiently rich for $(c^*, p^*) \equiv \left(c(\theta), p(\theta)\right)$ to be the unique maximizer of (4). for the highest match score $\bar{\theta}$. It is then straightforward to see that the profits of any firm that is assigned the sponsored slot are bounded by $(p^* - c^*)\left(1 - G(\bar{r} + p^*|\bar{\theta})\right)$.

We now propose an algorithm which is reminiscent of the platform's algorithm in our main analysis, and allows the platform to earn profits arbitrarily close to the upper bound for large enough n. The platform runs an adjusted second-price auction and places the firm with the highest bid in the sponsored slot and uniformly obfuscates all organic slots. Ties are broken in favor of higher match scores.

Proposition 4. *If firms differ in production costs and prices, there is a sequence of uniformly obfuscating algorithms that is asymptotically optimal. The expected match score of the firm that is placed in the sponsored slot converges to the highest possible match score $\bar{\theta}$.*

The idea behind Proposition 4 is similar to that underlying Theorem 1. As $n \to \infty$, the probability that a firm with cost-price pair $(c, p) \neq (c^*, p^*)$ wins the auction approaches 0. Nonetheless, even firms that do not have the optimal cost-price pair have an incentive to bid *as if* they had the optimal cost-price pair because there is a positive (but vanishing in the limit) probability that they value the sponsored slot the most because no firm with a more profitable cost-price pair exists. If a firm deviates downward, they earn zero profits in the limit, while an upward deviation to win the auction becomes too costly because firms with (c^*, p^*) almost surely exist so that outbidding them would imply negative profits. As a consequence, the auction for the sponsored slot is almost surely decided by the match score as a tiebreaker among firms with the cost-price pair (c^*, p^*). Since all these firms post equal prices and earn equal profits, the adjusted-second price auction underlying Theorem 1 remains asymptotically optimally when firms differ in their costs and prices.

4 Consumer Surplus

This section discusses the important policy question whether consumers are better or worse off if search platforms (stop) employ(ing) sponsored positions. We show that the answer to this question crucially depends on what other objectives a platform may have, besides revenue from selling a sponsored position. In particular, we consider two potential alternative objectives, and analyze the effects of introducing (or abandoning) a sponsored slot for each. First, a platform may directly care about consumer welfare (potentially as a short-hand for ad revenue). Second, as considered in the previous section, a platform may accrue revenue through sales commissions it receives when consumers buy products from firms.

The following two subsections deal with these two cases in turn, under the assumption that firms have identical prices. The last subsection considers the extent to which these results extend to firms having different prices.

4.1 Maximizing Consumer Welfare

We first focus on the case where in the absence of sponsored positions, the platform's and consumers' interests are aligned and the platform's objective is to maximize consumer surplus (which is maybe what Brin and Page had in mind when they wrote the phrase that we quoted in the Introduction). Common sense may suggest that the platform will always choose a perfect ranking, i.e., a ranking where firms with higher match score are ranked above firms with lower match scores. The next result says that this intuition is true if the number of firms to be allocated is arbitrarily large.

Proposition 5. *If a platform allocates firms to slots to maximize consumer surplus, then it is asymptotically optimal to choose a perfect ranking.*

When the number of firms n is small, the result may fail to hold as a perfect ranking may not give the consumer the best information possible of whether or not it is optimal to continue to search.[14] Essentially, if the platform has information that is contained in the match scores that is not perfectly transmitted to the consumer via a ranking, then providing a perfect ranking may be suboptimal even if preferences are aligned. If n grows large, the concern that a consumer may search for too long diminishes as the chance that there are not enough firms that are worth inspecting becomes small. It follows that consumers have an

[14] An example for small n where the result fails to hold is available upon request.

incentive to follow a perfect ranking if n is large and that the platform has an incentive to provide such a ranking. In fact, when n is large a perfect ranking delivers two advantages to consumers. First, the expected number of searches needed to find a satisfactory product is smaller than under a less informative ranking. Second, consumers more quickly learn when it is optimal to abort their search altogether as the realized match values are informative about when the chance that they will find a valuable object diminishes.

Corollary 2 compares the consumer-optimal ranking with our main result.

Corollary 2. *If, in the absence of a sponsored slot, the platform maximizes consumer welfare, then introducing a sponsored slot reduces consumers welfare if n is large.*

Note that Corollary 2 directly follows from Blackwell informativeness as consumers could always get the same pay-off as under uniform obfuscation by randomizing their search themselves. What is interesting is that the change in objectives does not affect the allocation of the top slot. Under both objectives this slot is allocated to the firm with the highest match score. It is rather the allocation of the remainder of firms across organic slots that lowers consumer welfare when the platform maximizes its revenue from the sponsored slot. Moreover, the platform's optimal ranking is far from obvious if its objective considers consumer welfare and auction revenue at the same time. Nevertheless, it remains true that the consumer is weakly better off without a sponsored slot.

4.2 Revenues from Sales Commissions

We next turn to the case where the platform earns a commission fee when the consumer buys from any of the firms in its list. As argued in Subsection 3.1, in the absence of a sponsored slot, the platform has an incentive to uniformly obfuscate all slots. With a sponsored slot, however, the sponsored slot contains relevant information to the consumer. From the increased informativeness of the ranking, it immediately follows that introducing a sponsored slot benefits consumers.

Corollary 3. *For large enough values of n if, in the absence of a sponsored slot, the platform maximizes sales commission revenue, then introducing a sponsored slot increases consumer welfare.*

Intuitively, the additional information provided to the consumer via the sponsored slot has two effects. First, it ensures that the consumer samples the firm that is most likely to have a high match value. This raises the expected match value of the product the consumer eventually chooses. Second, sampling the best match reduces the number of slots the consumer expects to inspect, thereby lowering expected search costs.

4.3 Surplus and Heterogeneous Prices

The previous two subsections have shown that in case all firms have the same prices (and margins) a sponsored slot increases consumer welfare if the platform would otherwise maximize sales commissions and decreases consumer welfare if the platform would otherwise maximize consumer welfare (as a short-hand for ad revenue). In this subsection we allow firms to have different prices and in that case, it is really the price-cost margin that is important to understand how equilibrium outcomes relate to consumer welfare.

First, consider that firms with higher prices have lower margins. This may be the case if production cost is an important determinant of prices and firms are not able or willing to fully pass on higher cost. In that case, as firms with higher price-cost margins have lower prices, it is not difficult to see that the results of the previous two subsections continue to hold as is implied by our next result, which focuses specifically on the content of the sponsored slot.

Proposition 6. *If firms with higher mark-ups post lower prices, i.e.,* $p - c \geq p' - c' \Leftrightarrow p \leq p'$, *then the sequence of uniformly obfuscating algorithms that is asymptotically optimal for the platform also maximizes expected consumer utility from inspecting the sponsored slot only.*

Intuitively, firms with higher price-cost margins are willing to bid more for the sponsored slot than firms with lower margins. When n is large there are, in expectation, enough firms with the highest price-cost margins. Consequently, one of them will be allocated the sponsored position, implying that the sponsored firm is also the firm that has the lowest price. In this case there is no conflict of interest between firms and consumers. It follows, that (as in the previous two subsections) having a sponsored slot is better for consumers if the platform otherwise maximizes commission fees, but not if it otherwise maximizes consumer welfare.

However, it need not always be the case that firms with the lowest prices have the highest mark-ups. Then, an asymptotically optimal sequence of uniformly obfuscating algorithms does not necessarily maximize expected consumer utility from inspecting the sponsored slot. To see why, let $p^s = \min_{h_p(p)>0} p \in \mathcal{P}$. By Lemma C.2, the sponsored slot goes to a firm with the lowest price p^s and the associated cost c^s if and only if (c^s, p^s) satisfies

$$(p^s - c^s)\left(1 - G(\bar{r} + p^s|\bar{\theta})\right) \geq (p - c)\left(1 - G(\bar{r} + p|\bar{\theta}\right)$$

for all $(c, p) \in C \times \mathcal{P}$ with $h(c, p) > 0$. Clearly, there is a $\varepsilon > 0$ and $\delta > 0$ such that

$$((p^s + \varepsilon) - (c^s - \delta))\left(1 - G(\bar{r} + (p^s + \varepsilon)|\bar{\theta})\right) > (p^s - c^s)\left(1 - G(\bar{r} + p^s|\bar{\theta})\right).$$

Hence, if there is a pair $(c, p) = (c^s - \delta, p^s + \varepsilon) \in C \times \mathcal{P} \setminus \{(c^s, p^s)\}$, then consumers do not find the firm with the lowest price in the sponsored slot. Since for large enough n, the highest match score of firms that charge p_s converges to $\bar{\theta}$ as well, consumer utility from inspecting the sponsored slot would not be maximized in those cases.

Nonetheless, even in this case the platform would not want to put firms that generate a very low surplus for the consumer in the sponsored position. Recall that if $n \to \infty$, there are arbitrarily many organic slots with a reservation value of $\bar{r}$. Thus, if the value of opening the sponsored slot falls short of $\bar{r}$, then consumer will never start their search at the sponsored slot, and the probability of ever inspecting the sponsored slot converges to zero in the limit. In this case, firms would not prefer the the sponsored slot over an organic slot, implying zero profits for the platform. Thus, the platform must select firms with a cost-price pair so that the value of opening the sponsored slot is at least $\bar{r}$. That is, $r_s \leq \bar{r}$ has to hold, where r_s is the reservation value of the sponsored slot implicitly defined in (**??**).

5 Discussion and Conclusion

In this paper, we have analyzed how selling sponsored positions affects a search platform's ranking of products. When deciding on its ranking, the platform takes into account that consumers are free to choose how to search. The platform has an incentive to use its information about consumer preferences to put the firm that it considers best matches the consumer preferences in the sponsored slot. This gives consumers an incentive to start their search there (which increases the sponsored firm's demand) and implies that winning the sponsored slot is positive news to firms. Obfuscation of organic slots also plays a crucial role as it increases the incentive for firms to acquire the sponsored slot by lowering the consumer's pay-off of searching beyond the sponsored slot, thereby boosting the sales (and revenues) from having the sponsored position.

Importantly, these results apply when the number of firms selling via the platform is sufficiently large as is true in many instances where consumers look for products on search platforms using keyword queries. With a small number of keyword relevant firms, learning effects arise and the optimal platform rule depends on specific details of the environment. We have demonstrated the robustness of our main result to (i) the platform also earning sales commissions, (ii) firms holding private information about the match value of their product and (iii) firms having different prices and costs.

The result has important implications for consumer welfare. We find that introducing a sponsored position harms consumers if otherwise the platform's maximizes consumer utility. This is because with sponsored positions, the platform has an incentive to reduce the informational value of the organic slots. If, however, in the absence of sponsored positions, the platform maximizes sales commissions revenue–as is common for many search platforms–then introducing a sponsored slot benefits consumers as it induces a strictly more informative ranking. If prices and price-cost margins differ across firms, these welfare results also depend on firms with lower prices having higher margins. But even when this is not the case, the platform is constrained by how consumers search. If platforms consistently allocate sponsored positions to firms that consumers do not appreciate, they will not inspect these slots (first). This significantly lowers the incentives for firms to bid for these slots and thereby inhibits the earning potential of these "preferred access" positions.

The paper focuses on the important question whether the large data sets on consumer search and purchase behavior that search platforms possess are used to benefit or harm consumers. Building on our paper, we see several fruitful directions for future research. For example, throughout this paper we have restricted the platform to sell a single sponsored position. Many platforms have multiple sponsored positions, however, and it is important to know what is the optimal way to allocate these slots. As a first step towards such an analysis, it is not difficult to see that this paper's results can be generalized to a platform selling a fixed number k of sponsored positions. The platform can randomly allocate the k firms with the highest match scores to the k sponsored positions while fully obfuscating the organic positions. With exogenously given prices the platform cannot do better than that. However, a firm that conducts the majority of its business through a search platform may set its prices strategically if it knows it often ends up among the top spots. It is unclear how the platform's incentives change in response. Additionally, it is also important to understand what determines the optimal number of sponsored positions in the first place.

References

C. D. Aliprantis and K. C. Border. *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Springer Science & Business Media, 2006.

S. P. Anderson and R. Renault. Pricing, product diversity, and search costs: A Bertrand-Chamberlin-Diamond model. *The RAND Journal of Economics*, 30(4):719–735, 1999.

S. P. Anderson and R. Renault. Search Direction: Position Externalities and Position Auction Bias. *CEPR DP # 16724*, 2021.

M. Armstrong and J. Zhou. Consumer Information and the Limits to Competition. *American Economic Review*, 112(2):534–77, 2022.

S. Athey and G. Ellison. Position auctions with consumer search. *The Quarterly Journal of Economics*, 126(3):1213–1270, 2011.

E. J. Balder. Generalized equilibrium results for games with incomplete information. *Mathematics of Operations Research*, 13(2):265–276, 1988.

H. Bar-Isaac and S. Shelegia. Monetizing Steering. *CEPR DP #17685*, 2022.

S. Brin and L. Page. The anatomy of a large-scale hypertextual web search engine. *Computer Networks and ISDN Systems*, 30:107–117, 1998.

Y.-K. Che and J. Hörner. Recommender systems as mechanisms for social learning. *The Quarterly Journal of Economics*, 133(2):871–925, 2018.

Y. Chen and C. He. Paid placement: Advertising and search on the internet. *The Economic Journal*, 121(556):F309–F328, 2011.

M. Dewatripont, I. Jewitt, and J. Tirole. The Economics of Career Concerns, Part I: Comparing Snformation Structures. *The Review of Economic Studies*, 66(1):183–198, 1999.

R. Donnelly, A. Kanodia, and I. Morozov. Welfare Effects of Personalized Rankings. *SSRN # 3649342*, 2022.

K. Eliaz and R. Spiegler. A simple model of search engine pricing. *The Economic Journal*, 121(556):F329–F339, 2011.

D. Garcia and S. Shelegia. Consumer search with observational learning. *The RAND Journal of Economics*, 49(1):224–253, 2018.

A. Ghose, P. G. Ipeirotis, and B. Li. Examining the Impact of Ranking on Consumer Behavior and Search Engine Revenue. *Management Science*, 60(7):1632–1654, 2014.

J. Glazer, I. Kremer, and M. Perry. The wisdom of the crowd when acquiring information is costly. *Management Science*, 67(10):6443–6456, 2021.

M. Janssen and C. Williams. Influencing Search. *CEPR DP # 15811*, 2022.

M. C. W. Janssen, A. Parakhonyak, and A. Parakhonyak. Non-reservation price equilibria and consumer search. *Journal of Economic Theory*, 172:120–162, 2017.

E. Kamenica and M. Gentzkow. Bayesian persuasion. *American Economic Review*, 101(6): 2590–2615, 2011.

T. T. Ke, S. Lin, and M. Y. Lu. Information design of online platforms. *HKUST Business School RP # 2022-070*, 2022.

E. Lehmann. Comparing location experiments. *Annals of Statistics*, 16(2):521–533, 1988.

E. Mossel, M. Mueller-Frank, A. Sly, and O. Tamuz. Social learning equilibria. *Econometrica*, 88(3):1235–1267, 2020.

M. Motta and A. Penta. Market effects of sponsored search auctions. *BSE Working Paper 1356*, 2022.

V. Nocke and P. Rey. Consumer Search, Steering and Choice Overload. *Journal of Political Economy*, forthcoming, 2023.

N. Persico. Information acquisition in auctions. *Econometrica*, 68(1):135–148, 2000.

M. Rothschild. Searching for the Lowest Price when the Distribution of Prices is Unknown. *Journal of Political Economy*, 82(4):689–711, 1974.

R. M. Ursu. The power of rankings: Quantifying the effect of rankings on online consumer search and purchase decisions. *Marketing Science*, 37(4):530–552, 2018.

M. L. Weitzman. Optimal search for the best alternative. *Econometrica*, 47(3):641–654, 1979.

A. Wolinsky. True monopolistic competition as a result of imperfect information. *The Quarterly Journal of Economics*, 101(3):493–511, 1986.

A Appendix: Preliminaries

To prove our asymptotic results, we consider the framework with infinitely many firms $i \in \mathbb{N}$ and then embed the model with finitely many firms within it. Let $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots) \in \Theta$ be the vector of scores, $\mathbf{v} = (v_1, v_2, \dots) \in V$ the vector of match values, and z the nonatomic random variable with support Z that the platform uses as a randomization device. We maintain the distributional assumptions of Section 1, denote the probability measure on $\Theta \times V \times Z$ by μ and denote probabilities and expectations with respect to μ by $\mathbb{P}[\cdot]$ and $\mathbb{E}[\cdot]$, respectively. Letting $\mathbf{b} = (b_1, b_2, \dots) \in B$ denote the bid vector and X the set of firm permutations, the set of algorithms $\mathcal{A}$ constitutes the set of functions $a : B \times \Theta \times Z \to X$ such that $a(\mathbf{b}, \cdot)$ is measurable for all $\mathbf{b} \in B$. The set of payment rules $\mathcal{P}$ consists of the set of measurable functions $\rho : B \to B$ with $\rho = (\rho_1, \rho_2, \dots)$ such that $\rho_i(\mathbf{b}) \leq b_i$.

To embed the finite firm model into this framework, let $\mathcal{A}_n \subset \mathcal{A}$ be the subset of algorithms which only permute the first n firms, which is to say that $a(\mathbf{b}, \boldsymbol{\theta}, z) = x$ implies that $x(i) = i$ for all $i \geq n + 1$. The game with n firms is obtained by requiring the platform to select an algorithm in $\mathcal{A}_n$ and restricting the consumer to only be able to inspect the first n positions. After choosing the algorithm, what follows is a proper subgame played by the firms, the consumer, and nature. Let $\mathcal{X}_n \subset \mathcal{A}_n \times \mathcal{P}$ denote the nonempty subset of algorithms and payment rules for which an equilibrium of the subgame exists. For the remainder, assume some measurable selection of the set of equilibria of the subgames and define $\Pi : \bigcup_{n=1}^{\infty} \mathcal{X}_n \to \mathbb{R}$ to be the platform's expected profit and $D : \bigcup_{n=1}^{\infty} \mathcal{X}_n \to \mathbb{R}$ the demand for the sponsored firm given that the selected equilibria are played.

A.1 Consumer Search with Learning

We first present three lemmas that characterize the consumer's search problem in a general environment where the consumer learns while searching. We begin by establishing an algorithm-independent upper bound on the equilibrium expected match value acquired by a consumer from engaging in optimal search. The main feature of the model that produces these results is that firms which are known to have the largest possible match score $\bar{\theta}$. The reservation value for a firm known to have the highest possible match score is the unique value r^* that satisfies $d_i(r^*, \mu(\cdot|\theta_i = \bar{\theta})) = 0$.

Lemma A.1. *In every equilibrium of the game with n firms, the expected match value acquired by the consumer who has made a purchase after searching m firms is less than* $u^* \equiv \mathbb{E}[v_i|\theta_i = \bar{\theta}, v_i \geq r^*]$, *for all* $m \leq n$.

Proof. We first claim that the consumer immediately halts her search and buys from a firm upon inspecting its good and finding a match value above r^*. Upon inspecting a single firm, the expected utility from continuing to search depends on the consumer's beliefs about the match scores of the remaining firms. Informing the consumer precisely of the remaining firms' match scores yields a Blackwell improvement and thus must weakly increase the expected utility from continued search. Also, setting the remaining firms' match scores to each equal the largest possible score $\bar{\theta}$ weakly increases the expected utility from continued search. In this case, it is optimal to follow Weitzman's rule and thus the consumer will halt her search and buy from the first visited firm when its match value exceeds r^*. Furthermore, if the match values uncovered at the first $m - 1$ visited firms are below r^* and the consumer decides to visit an mth firm, then this same argument provides that the consumer will immediately buy from this mth firm whenever its value is above r^*. Thus, the claim inductively follows.

Suppose the consumer buys the good sold by firm i when its match score is θ_i and the consumer's expected utility from choosing optimally among the set of continuation strategies that exclude buying immediately from firm i is w. Then the consumer's expected match value in this event is $\mathbb{E}[v_i|\theta_i, v_i \geq w]$. Due to the fact that this expectation is increasing in w and θ_i as well our previous arguments ensuring that $w \leq r^*$, the expression is bounded above by $u^* \equiv \mathbb{E}[v_i|\theta_i = \bar{\theta}, v_i \geq r^*]$. Thus $u^* < \infty$ serves as an upper bound for the expected match value acquired by the consumer and is independent of the algorithm, equilibrium strategies, number of firms inspected by the consumer, and number of firms in the market. □

Using this upper bound, we can likewise bound the probability that a consumer engages in a lengthy search based on the fact that, if the search tends to last too long, the consumer's expected payoff must be negative.

Lemma A.2. *In every equilibrium of the game and at every decision node for the consumer, for each $\varepsilon > 0$ there exists an $m \in \mathbb{N}$ such that the probability that a consumer searches m or more additional firms is less than ε.*

Proof. Consider a decision node for the consumer in an equilibrium of the game and let h denote the consumer's information set. Let M denote the random variable equal to the number of additional firms inspected by a consumer. From the previous lemma, a consumer's expected match value given that she searches m' more firms is less than u^* for all $m' \in \mathbb{N}$. Hence, a consumer's expected utility is bounded above by

$$\sum_{m' \in \mathbb{N}} (u^* - m' \cdot s)\mathbb{P}(M = m'|h).$$

For a given $m \in \mathbb{N}$, the above expression is less than or equal to

$$\mathbb{P}(M \leq m|h)u^* + \mathbb{P}(M > m|h)(u^* - m \cdot s).$$

As the consumer's expected utility from optimal search must be nonnegative, the above expression must likewise be nonnegative, implying

$$\mathbb{P}(M > m|h) \leq \frac{u^*}{m \cdot s}.$$

Thus, regardless of the algorithm or number of firms, if $m > \frac{u^*}{\varepsilon \cdot s}$, then the probability that the consumer searches beyond m more firms is less than ε. □

Define an "m-restricted" consumer as the consumer in our model, except that we restrict the consumer to only be able to search at most m positions. Let U denote the expected utility for the "unrestricted" consumer in our model and let U_m denote the expected utility for the m-restricted consumer. We obtain the following bound on expected utility.

Lemma A.3. *Consider a decision node at which the consumer has visited k firms and let h denote the consumer's information set. For a given* $\varepsilon > 0$*, let m be large enough so that the probability of searching* $m - k$ *more firms is less than* ε*. Then* $\mathbb{E}[U|h] \leq \mathbb{E}[U_m|h] + \varepsilon \cdot u^*$.

Proof. At the specified decision node, let M denote the random variable corresponding to the number of additional firms searched by the consumer. The m-restricted consumer can always mimic the unrestricted consumer's strategy for the next $m - k$ steps of search and then exit the market whenever the unrestricted consumer would search further, implying that $\mathbb{E}[U_m|h] \geq \mathbb{E}[U|M \leq m-k, h]\, \mathbb{P}(M \leq m-k, h)$. Therefore, we obtain

$$\begin{aligned}\mathbb{E}[U|h] &= \mathbb{E}[U|M \leq m-k, h]\, \mathbb{P}(M \leq m-k|h) + \mathbb{E}[U|M > m-k, h]\, \mathbb{P}(M > m-k|h) \\ &\leq \mathbb{E}[U_m|h] + \mathbb{E}[U|M > m-k, h]\, \mathbb{P}(M > m-k|h) \leq \mathbb{E}[U_m|h] + \varepsilon \cdot u^*.\end{aligned}$$

□

A.2 Mixing

A key idea in our argument is that when there are many firms, then for *any* possible algorithm, when the consumer finds that his match value with the sponsored firm lies below $\bar{r}$, he almost certainly will continue to search than to buy the sponsored firm's product. The tool we use to make this simple idea concrete is the fact that independently and identically distributed

(IID) random variables have the property of *mixing* (Mossel, Mueller-Frank, Sly, and Tamuz, 2020, Lemma 1). Intuitively, mixing means that any event E defined on the same probability space of a sequence of IID random variables $\{Y_i\}_{i\in\mathbb{N}}$ can only be strongly related to a finite number of them. Formally, for every $\epsilon > 0$, except for a set $N \subset \mathbb{N}$ with $|N| < 1/\epsilon^2$, each $i \notin N$ has the property that for every event K only depending on Y_i

$$\Big|\mathbb{P}(E \cap K) - \mathbb{P}(E)\,\mathbb{P}(K)\Big| < \epsilon. \tag{5}$$

Random variables satisfying (5) are called *ϵ-independent of E*. In our model, match values $\{v_i\}_{i\in\mathbb{N}}$ form an IID sequence of random variables and hence have the mixing property.

Proof of Lemma 1. For a given v and neighborhood N of $d_i(v,\mu)$, there exists an $\epsilon_0 > 0$ such that, if $|\lambda(K) - \mu(K)|$ for all $\sigma(v_i)$-measurable $K \subset \Omega$, then $d_i(v,\lambda) \in N$. The conclusion follows by noting that, to the contrary, if the conclusion does not hold, then there is a sequence of probability measures $\{\lambda_k\}_{k\in\mathbb{N}}$ satisfying $|\lambda_k(K) - \mu(K)| < 1/k$ for all $\sigma(v_i)$-measurable K and $d_i(v,\lambda_k) \notin N$ for all $k \in \mathbb{N}$. But this implies a contradiction because the induced probability distribution over v_i given by λ_k weakly converges to the induced probability distribution over v_i given by μ and thus $d_i(v,\lambda_k) \longrightarrow d_i(v,\mu)$ (see Aliprantis and Border, 2006, Theorem 15.3).

Let $\epsilon = \alpha \cdot \epsilon_0$ and consider $n > 1/\epsilon^2$. From the mixing property (Mossel et al., 2020, Lemma 1), at least one firm $i \le n$ has a score that is ϵ-independent of E_n. It follows that for any $\sigma(v_i)$-measurable event K

$$\Big|\mathbb{P}(K|E_n) - \mathbb{P}(K)\Big| < \frac{\epsilon}{\mathbb{P}(E_n)} < \frac{\epsilon}{\alpha} = \epsilon_0 \tag{6}$$

and thus $d_i(v,\mu(\cdot|E_n)) \in N$. □

Recall that, during search, the consumer's strategy involves choosing positions in the ranking to search. Let $\hat{v}_i = \hat{v}_i(\omega, a_n, \rho_n, \sigma_n)$ be the match value of the firm located in position i and define the function $\hat{d}_i(v,\lambda) = s - \int_\Omega \max\{\hat{v}_i - v, 0\}d\lambda(\omega)$. Under the hypotheses of Lemma 1, the conclusion extends to ensure the existence of an n^{**} such that $n \ge n^{**}$ implies that $|\hat{d}_i(v,\mu(\cdot|E_n)) - d(v,\mu)| < \delta$ for some position $i \le n$. This can be seen by first letting ϵ be the value which guarantees $|d_i(v,\mu(\cdot|E_n)) - d_i(v,\mu))| < \delta/2$ whenever v_i is ϵ-independent of E_n. The conclusion follows from noting that at least one position has probability of at least $1 - \frac{1}{\epsilon^2 n}$ of containing a match value that is ϵ-independent of E_n and then applying the triangle inequality.

B Appendix: Proofs for Section 2

Lemma B.1 (Upper Bound on Platform Profit)**.** *For any sequence of algorithms, payment rules, and corresponding selected equilibria,* $\{(a_n, \rho_n, \sigma_n)\}_{n\in\mathbb{N}}$*, the platform's limiting profit is bounded above by* $\limsup_{n\to\infty} \Pi(a_n, \rho_n) \leq p\left(1 - G(\bar{r}|\bar{\theta})\right)$.

Proof. First, we claim that $\limsup_{n\to} D(a_n, \rho_n) \leq 1 - G(\bar{r}|\bar{\theta})$. Notice that if $S_n \subset \Theta\times V\times Z$ is the event in which the consumer buys from the sponsored firm in equilibrium σ_n, for a given $\delta > 0$ we have

$$D(a_n, \rho_n) = \mathbb{P}(S_n \cap \{\hat{v}_1 \geq \bar{r} - \delta\}) + \mathbb{P}(S_n \cap \{\hat{v}_1 < \bar{r} - \delta\}). \tag{7}$$

Denoting $E_n \equiv S_n \cap \{\hat{v}_1 < \bar{r} - \delta\}$, we now show that $\mathbb{P}(E_n) \to 0$ as $n \to \infty$. Toward a contradiction, suppose that $\mathbb{P}(E_n)$ does not vanish in the limit, which is to say that there is an $\alpha > 0$ and a subsequence $\{n_k\}_{k\in\mathbb{N}}$ satisfying $\mathbb{P}(E_{n_k}) > \alpha$ for all k. From Lemma 1, there is a k^* such that, for all $k \geq k^*$, there is a position $i \leq n_k$ with $\hat{d}_i(\hat{v}_i, \mu(\cdot|E_{n_k})) < \hat{d}_i(\bar{r} - \delta, \mu(\cdot|E_{n_k})) < 0$. But this implies a contradiction as the consumer can yield a strictly higher expected payoff by deviating and committing to inspect (or possibly paying the search cost to reinspect) position i whenever she knows that the event E_{n_k} has occurred. Therefore $\mathbb{P}(E_n)$ must vanish in the limit.

Next, observe that

$$\mathbb{P}(S_n \cap \{\hat{v}_1 \geq \bar{r} - \delta\}) \leq \mathbb{P}(\hat{v}_1 \geq \bar{r} - \delta) \leq 1 - G(\bar{r} - \delta|\bar{\theta}).$$

Thus, it follows that $\limsup_{n\to\infty} D(a_n, \rho_n) \leq 1 - G(\bar{r} - \delta|\bar{\theta})$. Moreover, as the inequality holds for all $\delta > 0$, we have $\limsup_{n\to\infty} D(a_n, \rho_n) \leq 1 - G(\bar{r}|\bar{\theta})$. This proves our first claim.

Using this bound placed on the limiting demand, we now bound the platform's limiting profit. The platform's profit is less than or equal to the winning bid. Let β_n denote the expected winning bid in equilibrium σ_n. Whenever a firm has a positive probability of winning the auction, its equilibrium bid cannot exceed the expected profit conditional on winning. That is, its bid is bounded above by $p \cdot D(a_n, \rho_n)$. The limiting equilibrium bids, therefore, satisfy $\limsup_{n\to\infty} \beta_n \leq \limsup_{n\to\infty} p \cdot D(a_n, \rho_n) \leq p(1 - G(\bar{r}|\bar{\theta}))$. □

Using uniform obfuscation, it becomes possible to approximately guarantee that the consumer buys from the sponsored firm whenever its match value exceeds $\bar{r}$.

Lemma B.2 (Obfuscation Achieves Bound)**.** *Let* $\{(a_n, \rho_n, \sigma_n)\}_{n\in\mathbb{N}}$ *be a sequence of uniformly obfuscating algorithms, payment rules, and selected equilibrium and suppose that for all sufficiently large* n*, the consumer begins search at the sponsored position. Then the probability that the consumer receives a match value above* $\bar{r}$ *and searches at least one organic firm vanishes in the limit.*

Proof. Let $\delta > 0$ and let E_n denote the event in which the consumer receives a match value above $\bar{r} + \delta$ from the sponsored firm and searches at least one organic firm. We begin by proving the claim for the simpler problem where we replace the consumer in our model with an "m-restricted consumer" who shares the same preferences, but can search at most m firms. Combining this conclusion with Lemma A.3 then proves the original claim.

Fix $m \geq 2$ and let $E_{m,n}$ denote the event in which the m-restricted consumer searches m firms and $\hat{v}_1 \geq \bar{r} + \delta$. We claim that $\lim_{n\to\infty} \mathbb{P}(E_{m,n}) = 0$. Toward a contradiction, suppose to the contrary that there is a constant $\alpha > 0$ and a subsequence $\{n_k\}_{k=1}^{\infty}$ along which $\mathbb{P}(E_{m,n_k}) > \alpha$ holds for all k. By Lemma 1, there is an n^* for which $n \geq n^*$ implies that there is some position $i \leq n$ satisfying $\hat{d}_i(\hat{v}_1, \mu(\cdot|E_{n_k})) > \hat{d}_i(\bar{r} + \delta, \mu(\cdot|E_{n_k})) > 0$. Because the algorithm is uniformly obfuscating, the match value distributions are the the same for all $i \in \{m, m+1, \ldots, n\}$ and thus $\hat{d}_m(\bar{r} + \delta, \mu(\cdot|E_{n_k})) > 0$. This means that buying from the sponsored position is, in expectation, strictly preferred to inspecting the mth position conditional on the event E_{m,n_k}. But this implies a contradiction as E_{m,n_k} is an event in which the consumer inspects the mth position. Therefore, $\lim_{n\to\infty} \mathbb{P}(E_{m,n}) = 0$.

Similarly, consider the case with $m \geq 3$ and let $E_{(m-1),n}$ be the event in which $\hat{v}_1 \geq \bar{r} + \delta$ and the m-restricted consumer searches at least $m - 1$ firms. Given $E_{(m-1),n}$, the expected difference in the expected utility between buying the good with the highest match value of those which have been inspected in the first $m - 2$ positions and continuing search to the $m - 1$th position is greater than[15]

$$\hat{d}_{m-1}(\bar{r} + \delta, \mu(\cdot|E_{(m-1),n})) - \mathbb{P}(E_{m,n}|E_{(m-1),n}) \cdot u^*.$$

By the same argument above, if $\mathbb{P}(E_{(m-1),n})$ has a subsequence that is bounded away

[15] Let $\mathbf{v}_k \equiv (\hat{v}_1, \ldots, \hat{v}_k)$. In the event $E_{(m-1),n}$, the realized utility is $(\max \mathbf{v}_{m-1} - (m-1) \cdot s)\mathbf{1}_{E_{m,n}^C} + (\max \mathbf{v}_m - m \cdot s)\mathbf{1}_{E_{m,n}} \leq \max \mathbf{v}_{m-1} - (m-1)s + \max \mathbf{v}_m \cdot \mathbf{1}_{E_{m,n}}$. Subtracting the right side from the realized utility if the consumer were to only inspect $m-2$ firms yields $s - \max\{\hat{v}_{m-1} - \max \mathbf{v}_{m-2}, 0\} - \max \mathbf{v}_m \cdot \mathbf{1}_{E_{m,n}}$. Taking the expectation of this expression given $E_{(m-1),n}$ and applying Lemma A.1 obtains $\hat{d}_{m-1}\left(\max \mathbf{v}_{m-2}, \mu(\cdot|E_{(m-1),n})\right) - \mathbb{E}[\max \mathbf{v}_m \cdot \mathbf{1}_{E_{m,n}}|E_{(m-1),n}] > \hat{d}_{m-1}(\bar{r} + \delta, \mu(\cdot|E_{(m-1),n})) - \mathbb{P}(E_{m,n}|E_{(m-1),n}) \cdot u^*$.

from zero, then this expression is eventually positive along the subsequence, implying a contradiction. Thus, $\mathbb{P}(E_{(m-1),n}) \to 0$. Continuing the argument, it inductively follows that $\mathbb{P}(E_{2,n}) \to 0$ for an m-restricted consumer with $m \geq 2$.

Let $0 < \epsilon < u^*/d_i(\bar{r} + \delta, \mu)$ and let m be large enough so that the probability of searching $m - 1$ more positions from a decision node is less than ϵ, as provided by Lemma A.2. Using Lemma A.3, conditional on the event E_n, the expected difference in the expected utility between buying the good from the sponsored position and searching the first organic position is greater than

$$\hat{d}_2\left(\bar{r} + \delta, \mu(\cdot|E_n)\right) - \mathbb{P}(E_{2,n}|E_n) \cdot u^* - \epsilon \cdot u^*. \tag{8}$$

Repeating the above arguments again finds that, if $\mathbb{P}(E_n)$ has a subsequence bounded away from zero, then (8) is eventually positive for a large enough n, implying a contradiction. Therefore, $\mathbb{P}(E_n) \to 0$. As the choice of $\delta > 0$ was arbitrary, the desired conclusion holds. □

Our interest is in equilibria in which all firms submit the same bid and the platform awards the sponsored position to the firm with the highest match score and performs uniform obfuscation within the organic positions. It is a straightforward calculation to show that, for such algorithms and bidding strategies, the consumer optimal begins his search at the sponsored position. The calculation can be found in §S.1 of the Supplementary Material. For these equilibria, the demand for a firm with score that wins the sponsored position and has a score of θ_i eventually exceeds $1 - G(\bar{r} + \delta|\theta_i)$ for all $\delta > 0$. Moreover, the distribution of the winning firm's score weakly converges to a point mass on $\bar{\theta}$, thus, the sequence of the uniformly obfuscating algorithms achieves the upper bound given in Lemma B.1.

Notice that if a firm deviates to a higher bid then it wins the auction for sure, but loses information about whether it has the highest match score. It must be that such a deviation offers an expected profit of $p(1 - G(\bar{r}))$ in the limit.

Lemma B.3. *Consider a strategy profile in which all firms place the same bid, the platform awards the sponsored position to the firm with the highest bid and breaks indifference in favor of the firm with the highest match score, and the consumer searches optimally given this strategy. Let S'_n be the event in which a firm which deviates to a higher bid makes a sale. Then* $\lim_{n\to+\infty} \mathbb{P}(S'_n) \leq 1 - G(\bar{r})$.

Proof. Let $S'_n \subset \Theta \times V \times Z$ and $S'_{m,n} \subset \Theta \times V \times Z$ be the events in which the consumer and m-restricted consumer buy the sponsored product given a deviation, respectively. As before, given the proposed strategies, it without loss of generality to assume that the consumer searches organic firms in order of their ranking. As the continuation value for the unrestricted consumer is always at least that of the m-restricted consumer, we have $S'_n \subset S'_{m,n}$ for all m and n. For a given $k \in \mathbb{N}$, let $\underline{v}_k \in (\underline{v}, \bar{r})$ satisfy $\frac{G(\bar{r}|\theta_i) - G(\underline{v}_k|\theta_i)}{G(\bar{r}|\theta_i)} \geq 1 - \frac{1}{k}$ for all θ_i. Let V^k denote the event in which $v_S \in [\underline{v}_k, \bar{r}]$ given the deviation. Having assumed that the density $g(v_i|\theta_i)$ is positive on the interior and continuous, we know that there is a constant $c > 0$ such that $\frac{g(v_i|\theta_i)}{g(v_i|\theta_j)} \geq c$ whenever $v_i \in V^k$ for all θ_i and θ_j.

Given the bound placed on the likelihood ratio in the event V^k, the m-restricted consumer's beliefs about the first $m-1$ organic firm's match values converges to the true distribution, which is that they are IID $G(v_i)$. At the limiting distribution, the probability that the m-restricted consumer buys from the sponsored firm given V^k is $\int_{V^k} G(v_S)^{m-1} dG(v_S|V^k) < G(\bar{r})^{m-1}$. Using the inequality

$$\begin{aligned}\mathbb{P}(S'_{m,n}|v_S \leq \bar{r}) &\leq \mathbb{P}(S'_{m,n}|V^k)\,\mathbb{P}(V^k|v_S \leq \bar{r}) + 1 - \mathbb{P}(V^k|v_S \leq \bar{r}) \\ &\leq \mathbb{P}(S'_{m,n}|V^k)\left(1 - \frac{1}{k}\right) + \frac{1}{k}\end{aligned}$$

we obtain

$$\lim_{n\to\infty} \mathbb{P}(S'_{m,n}|v_S \leq \bar{r}) \leq \lim_{n\to\infty} \mathbb{P}(S'_{m,n}|V^k)\left(1 - \frac{1}{k}\right) + \frac{1}{k} < G(\bar{r})^{m-1}\left(1 - \frac{1}{k}\right) + \frac{1}{k}.$$

As the above expression holds for all $k \in \mathbb{N}$, taking the limit as $k \to \infty$ we find $\lim_{n\to\infty} \mathbb{P}(S'_{m,n}|v_S \leq \bar{r}) \leq G(\bar{r})^{m-1}$. Because $S'_n \subset S'_{m,n}$, it follows that $\lim_{n\to\infty} \mathbb{P}(S'_n|v_S \leq \bar{r}) \leq G(\bar{r})^{m-1}$ for all $m \in \mathbb{N}$ and therefore $\lim_{n\to\infty} \mathbb{P}(S'_n|v_S \leq \bar{r}) = 0$. The desired conclusion therefore follows from noting that $\mathbb{P}(S'_n) = \mathbb{P}(S'_n|v_S \leq \bar{r})G(\bar{r}) + \mathbb{P}(S'_n|v_S > \bar{r})(1 - G(\bar{r}))$. □

Proof of Theorem 1. Let $a^*_n \in \mathcal{A}_n$ be an algorithm which, for a given β_n: (1) if the highest bid exceeds β_n, assign a firm with the highest bid to the sponsored position, breaking ties in favor of a firm with the highest score and remaining ties with uniform probability; (2) if the highest bid is less than β_n, select a winner of the auction with uniform probability; and (3) assign those firms that do not win the auction to organic slots with uniform probability.

Let us provide the necessary conditions for there to be a symmetric equilibrium in which all firms submit the same bid equal to β_n. Suppose all firms bid β_n. A firm's expected profit from likewise bidding β_n is equal to

$$\frac{1}{n}(\pi(1, a_n) - \beta_n) + \sum_{m=2}^{n} \frac{1}{n}\pi(m, a_n) \tag{9}$$

where $\pi(m, a_n)$ is the firm's expected profit from playing the tentative equilibrium strategy and being relegated to the mth position. Deviating to a lower bid yields an expected profit of

$$\sum_{m=2}^{n} \frac{1}{n-1}\tilde{\pi}(m, a_n) \tag{10}$$

where $\tilde{\pi}(m, a_n)$ is a firm's expected profit from being in a position $m \geq 2$ following the deviation. Deviating to a higher bid yields an expected profit of

$$\hat{\pi}(1, a_n) - \beta_n \tag{11}$$

where $\hat{\pi}(1, a_n)$ denotes the firm's expected profit in the sponsored position following the upward deviation. Combining these conditions, we find bidding β_n to be a best reply if and only if

$$\frac{n}{n-1}\hat{\pi}(1, a_n) - \frac{1}{n-1}\sum_{m=1}^{n} \pi(m, a_n) \leq \beta_n \leq \sum_{m=1}^{n} \pi(m, a_n) - \frac{n}{n-1}\sum_{m=2}^{n} \tilde{\pi}(m, a_n). \tag{12}$$

Denote the leftmost side by λ_n and the rightmost side by ρ_n so that (12) simplifies to $\lambda_n \leq \beta_n \leq \rho_n$. We want to show that as the number of firms grows large $\lambda_n < \rho_n$.

Given the algorithm and the consumer's optimal search, we establish the following. First, because total industry profit is bounded from above, the product $\frac{1}{n-1}\sum_{m=1}^{n} \pi(m, a_n)$ vanishes in the limit. Second, by Lemma B.3, the expected profit when deviating to a higher bid converges to $\hat{\pi}(1, a_n) \to p(1 - G(\bar{r}))$ since the consumer will only make a purchase if his match with the sponsored firm exceeds $\bar{r}$ and the distribution of the match value of the upward deviating firm is $G(v_i) = \int_{\underline{\theta}}^{\bar{\theta}} G(v_i|\theta_i)dF(\theta_i)$. Third, if we let $\tilde{\pi}^*(m, a_n)$ denote the expected profit for a firm deviating to a lower bid and being assigned to a position $m \geq 2$ given that it has the highest score, we have $\tilde{\pi}(m, a_n) = \frac{1}{n}\tilde{\pi}^*(m, a_n) + \frac{n-1}{n}\pi(m, a_n)$. Hence, we can write $\rho_n = \pi(1, a_n) - \frac{1}{n-1}\sum_{m=2}^{n} \tilde{\pi}^*(m, a_n)$ where the bound on industry profit guarantees that the rightmost term vanishes in the limit. Combining these three observations employing Lemma B.1, we find that $\lambda_n \to p(1 - G(\bar{r}))$ and $\rho_n \to p(1 - G(\bar{r}|\bar{\theta})) > p(1 - G(\bar{r}))$, where the inequality holds because $G(v|\theta)$ satisfies the MLRP. Thus, there exists an n^* such that $n \geq n^*$ implies $\lambda_n \leq \rho_n$.

Define $\beta_n = \rho_n$ whenever $\lambda_n \leq \rho_n$ and $\beta_n = 2p$ otherwise. Whenever $\lambda_n > \rho_n$, it is clear that each firm's best reply is to bid zero. Whenever $\lambda_n \leq \rho_n$, there is an equilibrium in which all firms bid β_n. Therefore, we obtain $\limsup_{n\to\infty} \beta_n \leq \limsup_{n\to\infty} \Pi(a_n^*)$ and also $\limsup_{n\to\infty} \beta_n = p(1 - G(\bar{r}|\bar{\theta}))$. Combining this with Lemma B.1, we can conclude that $\{a_n^*\}_{n\in\mathbb{N}}$ is asymptotically optimal if $\{a_n^*\}_{n\in\mathbb{N}}$ exhibits uniform obfuscation. □

Proof of Proposition 1. As the maximal limiting profit under the information structures are $p \cdot (1 - G(\bar{r}|\bar{\theta}))$ and $p \cdot (1 - \tilde{G}(\bar{r}|\bar{\theta}))$ respectively, the result follows from showing $\tilde{G}(\bar{r}|\bar{\theta}) \leq G(\bar{r}|\bar{\theta})$. Drawing from the argument for Theorem 5.1 in Lehmann (1988), let $\{\alpha_m\}$ be a vanishing sequence of values in $(0, 1)$, $\{t_m\}$ the sequence satisfying $F(t_m|\bar{r}) = 1 - \alpha_m$, and $\{\tilde{t}_m\}$ the sequence satisfying $\tilde{F}(\tilde{t}_m|\bar{r}) = 1 - \alpha_m$. Due to the fact that $\tilde{F}$ is more accurate than F, we have $F(t_m|v_i) \leq \tilde{F}(\tilde{t}_m|v_i)$ for all $v_i < \bar{r}$ and $F(t_m|v_i) \geq \tilde{F}(\tilde{t}_m|v_i)$ for all $v_i > \bar{r}$. Consider the two posterior probabilities

$$G(\bar{r}|\theta_i \geq t_m) = \frac{\int_{\underline{v}}^{\bar{r}} (1 - F(t_m|v_i))\, g(v_i) dv}{\int_{\underline{v}}^{\bar{v}} (1 - F(t_m|v_i))\, g(v_i) dv}$$

$$\tilde{G}(\bar{r}|\theta_i \geq \tilde{t}_m) = \frac{\int_{\underline{v}}^{\bar{r}} \left(1 - \tilde{F}(\tilde{t}_m|v_i)\right) g(v_i) dv}{\int_{\underline{v}}^{\bar{v}} \left(1 - \tilde{F}(\tilde{t}_m|v_i)\right) g(v_i) dv}.$$

By rearranging terms, we see that $\tilde{G}(\bar{r}|\theta_i \geq \tilde{t}_m) \leq G(\bar{r}|\theta_i \geq t_m)$ if and only if

$$\frac{\int_{\underline{v}}^{\bar{r}} \left(1 - \tilde{F}(\tilde{t}_m|v_i)\right) g(v_i) dv}{\int_{\underline{v}}^{\bar{r}} (1 - F(t_m|v_i))\, g(v_i) dv} \leq \frac{\int_{\bar{r}}^{\bar{v}} \left(1 - \tilde{F}(\tilde{t}_m|v_i)\right) g(v_i) dv}{\int_{\bar{r}}^{\bar{v}} (1 - F(t_m|v_i))\, g(v_i) dv} \tag{13}$$

which must hold as the left side is less than one while the right side exceeds one. Thus, $\tilde{G}(\bar{r}|\theta_i \geq \tilde{t}_m) \leq G(\bar{r}|\theta_i \geq t_m)$ for all m. At the same time, $G(\bar{r}|\theta_i \geq t_m) \to G(\bar{r}|\bar{\theta})$ and $\tilde{G}(\bar{r}|\theta_i \geq \tilde{t}_m) \to \tilde{G}(\bar{r}|\bar{\theta})$ as $m \to +\infty$, implying $\tilde{G}(\bar{r}|\bar{\theta}) \leq G(\bar{r}|\bar{\theta})$. □

C Appendix: Proofs for Section 3

C.1 Sales Commissions

Proof of Proposition 2. The statement follows directly from the proof of Theorem 1 setting $\eta = \bar{r}$. □

C.2 Privately Informed Firms

Consider an algorithm which awards the sponsored slot to the firm with the highest adjusted bid and uniformly and randomly assigns all other firms to organic positions. Denote $\bar{b} \equiv p(1 - G(\bar{r}|\bar{t} + \bar{\theta}))$. A firm's **adjusted bid** $\psi(b_i, \theta_i)$ is a smooth and strictly increasing function of firm's bid on $[0, \bar{b}]$ and score on $[\underline{\theta}, \bar{\theta}]$. Normalize the smallest adjusted bid to $\psi(0, \underline{\theta}) = 0$. We suppose that for all bids $b_i > \bar{b}$ the adjusted score is the same as if they had bid zero $\psi(b_i, \theta_i) = \psi(0, \theta_i)$. Specify the payment rule so that the firm that wins the auction pays the value of its bid. As bidding higher than $\bar{b}$ is dominated by bidding zero, it is without loss in generality to restrict the firms' strategy space to bids in $[0, \bar{b}]$. Because the distribution of θ_i is atomless, the probability of tying is zero. Thus, each firm's expected profit is continuous in the bids on $[0, \bar{p}]$.

Using Lemma B.2, we pin down a lower bound on the expected profit for a firm from winning the auction when n grows large.

Lemma C.1. *For all $\epsilon' > 0$ and $\bar{v}' \in (\bar{r}, \bar{v})$ there exists $n' \in \mathbb{N}$ such that, if $n \geq n'$, then a firm with signal $y_i = (\theta_i, t_i)$ that wins the auction generates an expected profit that is greater than $p(G(\bar{v}'|y_i) - G(\bar{r}|y_i)) - \epsilon'$.*

Proof. Let $\mathcal{V}_n \subset [\bar{r}, \bar{v}']$ denote the subset of match values such that the consumer buys immediately from the sponsored firm if $\hat{v}_1 \in [\bar{r}, \bar{v}'] \backslash \mathcal{V}_n$ and searches at least one more firm $\hat{v}_1 \in \mathcal{V}_n$. From the Lemma B.2, the Lebesgue measure of $\mathcal{V}_n$ must vanish in the limit. Thus, when a firm with signal $y_i = (\theta_i, t_i)$ wins the auction, it makes a sale immediately with probability $G(\bar{v}'|y_i) - G(\bar{r}|y_i) - \int_{\mathcal{V}_n} dG(v_i|y_i)$.

We argue that $\int_{\mathcal{V}_n} dG(v_i|y_i)$ uniformly converges to zero as $n \to \infty$. The density $g(v_i|y_i)$ is finite for all $y_i \in [\underline{y}, \bar{y}]$ and $v_i \in [\bar{r}, \bar{v}']$. Due to compactness and continuity of the density, there must be a constant $c > 0$ such that $\max_{v_i \in [\bar{r}, \bar{v}'], y_i \in [\underline{y}, \bar{y}]} g(v_i|y_i) \leq c$. But then it follows that $\max_{y_i \in [\underline{y}, \bar{y}]} \int_{\mathcal{V}_n} g(v_i|y_i) dv_i \leq c \cdot \lambda(\mathcal{V}_n)$ and uniform convergence follows. Thus, it follows that, for all $\epsilon' > 0$ there exists $n' \in \mathbb{N}$ such that, if $n \geq n'$, then a deviating firm's profit given that it wins the auction and has a signal of y_i is greater than $p(G(\bar{v}'|y_i) - G(\bar{r}|y_i)) - \epsilon'$. □

Proof of Proposition 3. Consider the sequence of uniformly obfuscating algorithms described above. For the moment, assume that the consumer begins search at the sponsored position. The calculation found in Section S.1 verifies that this is optimal in the limit. As all payoffs are continuous in each player's strategy and the strategy spaces are compact, Theorem 3.1 in Balder (1988) establishes the existence of an equilibrium. The argument in Lemma B.1 provides that the limit inferior of the profit derived from any sequence of algorithms is bounded above by $\bar{b}$. We want to show that the uniformly obfuscating algorithms described above achieve this upper bound. Towards this end, we begin by proving the following claim. For each $n \geq 2$, let σ_n be an equilibrium strategy profile for the game featuring n firms.

Let Ψ_n be the random variable that corresponds to the highest adjusted bid in equilibrium σ_n. We claim that Ψ_n weakly converges to point mass on $\bar{\psi} \equiv \psi(\bar{b}, \bar{\theta})$. Toward a contradiction, suppose to the contrary that there is a $\psi^* < \bar{\psi}$, an $\epsilon > 0$, and a subsequence $\{n_k\}_{k \in \mathbb{N}}$ such that $\mathbb{P}(\Psi_{n_k} \leq \psi^*) \geq \epsilon$ for all k. Let i_n be a firm with the lowest expected profit in equilibrium σ_n. The expected profit for i_n must vanish in the limit. Let $\hat{\Psi}_n$ correspond to the highest adjusted bid among firms $j \neq i_n$. As $\hat{\Psi}_n \leq \Psi_n$, it must also be that $\mathbb{P}(\hat{\Psi}_{n_k} \leq \psi^*) \geq \epsilon$ for all k. Let $M = \{\lambda \in \Delta([0, \bar{\psi}]) : \lambda([0, \psi^*]) \geq \epsilon\}$. Given that $\Delta([0, \bar{\psi}])$ is compact and M is a closed subset of $\Delta([0, \bar{\psi}])$, M is also compact. We want to show that, for n large enough, firm i_n can achieve a positive expected payoff whenever the distribution of $\hat{\Psi}_n$ is given by a probability measure $\lambda \in M$.

Let b^* denote the bid that satisfies $\psi(b^*, \bar{\theta}) = \psi^*$. Let $\epsilon' > 0$ be small enough and $\bar{v}'$ large enough so that $p(G(\bar{v}'|\bar{t} + \bar{\theta}) - G(\bar{r}|\bar{t} + \bar{\theta})) - \epsilon' > b^*$. From Lemma C.1, there is an $n' \in \mathbb{N}$ such that $n \geq n'$ implies that the expected net profit from winning the auction for a firm i with signal t_i and score θ_i is at least $p(G(\bar{v}'|\theta_i + t_i) - G(\bar{r}|\theta_i + t_i)) - \epsilon'$.

Suppose that the distribution of $\hat{\Psi}_n$ is $\lambda \in M$. Notice that if $\lambda\left(\left[0, \psi\left(0, \bar{\theta}\right)\right)\right) > 0$, then firm i_n's expected profit from always bidding zero is positive. Suppose instead that $\lambda\left(\left[0, \psi\left(0, \bar{\theta}\right)\right)\right) = 0$ so that i_n's bid must be positive to ensure a positive chance of winning the auction. Let $\underline{\psi} = \min \operatorname{supp} \lambda$. Notice that if i_n submits the bid $\underline{b}$ that satisfies $\psi(\underline{b}, \bar{\theta}) = \underline{\psi}$, then the probability of i_n winning the auction is zero. All bids higher than $\underline{b}$ deliver a positive probability of winning the auction. Letting $b_{i_n} \to \underline{b}$ from the right, the distribution of i_n's score conditional on winning the auction weakly converges to a point mass on $\bar{\theta}$. When t_{i_n} is in a neighborhood of $\bar{t}$, then $\iint p(G(\bar{v}'|\theta_{i_n} + t_{i_n}) - G(\bar{r}|\theta_{i_n}, t_{i_n}))dF(\theta_{i_n}|\psi(b_{i_n}, \theta_{i_n}) \geq x)d\lambda(x) - \epsilon' > b^* > b_{i_n}$. This means that for all $n \geq n'$, if the distribution of $\hat{\Psi}_n$ is λ, then i_n has a positive expected profit. Since M is compact and firm i_n's expected profit is continuous in the adjusted bid distribution for all other firms, then firm i_n's expected profit is bounded away from zero for all $n \geq n'$. But this implies a contradiction and thus

$\liminf_{n\to\infty} \mathbb{P}(\Psi_n \geq \psi^*) = 1$ for all $\psi^* < \bar{\psi}$. The desired conclusion follows. □

C.3 Heterogeneous Prices

Asymptotically, the profits of the firm that is assigned the sponsored slot are bounded by $(p^* - c^*)\left(1 - G(\bar{r} + p^*|\bar{\theta})\right)$. It follows that this expression also constitutes a bound for platform profits. Moreover, due to the independence of prices and match scores/values, just as in our main analysis, uniform obfuscation implies that a consumer who continues to search beyond the sponsored slot never returns a.s. as $n \to \infty$.

Lemma C.2. *There exists an $\bar{n}$ large enough so that firms with the price cost pair (c^*, p^*) bid more in equilibrium than any firm with $(c, p) \in C \times \mathcal{P} \setminus (c^*, p^*)$ and $h(c, p) > 0$ if $n \geq \bar{n}$.*

Proof. Consider the expected difference in profits from obtaining and not obtaining the sponsored slot:

$$(p - c)\left(1 - G(\bar{r} + p|\hat{\theta}^{c,p}(n, H)\right) - \frac{1}{n-1}\sum_{m=2}^{n} \pi\left(m, c, p, \hat{\theta}^{c,p}(n, H)\right) \tag{14}$$

where $\hat{\theta}^{c,p}(n, H)$ denotes the expected highest match score among all firms with the cost price pair (c, p), given the distribution H and n and where $\pi(m, \cdot)$ is the profit from position m. Clearly, firms will never bid more than (14). This upper bound increases in θ. Thus, the most that a firm with $(c, p) \neq (c^*, p^*)$ bids is

$$(p - c)\left(1 - G(\bar{r} + p|\bar{\theta}\right) - \frac{1}{n-1}\sum_{m=2}^{n} \pi\left(m, c, p, \bar{\theta}\right). \tag{15}$$

Note that the second term vanishes in the limit because the joint profits of all firms ares bounded. Then, by the properties of $G(\cdot)$ there is a $\theta^* < \bar{\theta}$ such that

$$(p - c)\,(1 - G(\bar{r} + p|\theta)) < (p^* - c^*)\,(1 - G(\bar{r} + p^*|\theta')) \tag{16}$$

for all $(c, p) \in C \times \mathcal{P} \setminus (c^*, p^*)$ with $h(c, p) > 0$, $\theta \in \mathbf{\Theta}$ and $\theta' > \theta^*$. This is true since, by assumption

$$(p - c)\left(1 - G(\bar{r} + p|\bar{\theta})\right) < (p^* - c^*)\left(1 - G(\bar{r} + p^*|\bar{\theta})\right) \tag{17}$$

for all $(c, p) \in C \times \mathcal{P} \setminus (c^*, p^*)$ with $h(c, p) > 0$, and because $(p - c)\,(1 - G(\bar{r} + p|\theta))$ is continuous in θ, and $C \times \mathcal{P}$ is finite, there is a $\theta^* < \bar{\theta}$ such that

$$(p - c)\left(1 - G(\bar{r} + p|\bar{\theta})\right) < (p^* - c^*)\left(1 - G(\bar{r} + p^*|\theta')\right) \tag{18}$$

for all $(c, p) \in C \times \mathcal{P} \setminus (c^*, p^*)$ with $h(c, p) > 0$, $\theta \in \mathbf{\Theta}$ and $\theta' > \theta^*$. Since the left-hand side of the inequality above is monotonically increasing in θ, inequality (16) follows.

Since $\hat{\theta}^{c^*,p^*}(n, H)$ increases in n, there is an $\bar{n}$ large enough so that $\hat{\theta}^{c^*,p^*}(n, H) > \theta^*$. Considering (14) and again noting that the second term vanishes in the limit, we conclude that the upper bound of what (c^*, p^*) types are willing to bid exceeds that of any other type if $n \geq \bar{n}$. Clearly, firms with (c^*, p^*) will never bid less than any firm with $(c, p) \neq (c^*, p^*)$ in any symmetric equilibrium because if that were the case, then every firm with type (c^*, p^*) would find it profitable to raise its bid. □

Proof of Proposition 4. As $n \to \infty$, it is straightforward to see that the probability that there are at least two firms with cost-price pair (c^*, p^*) converges to 1. By Lemma C.2, these firms with (c^*, p^*) compete only against each other asymptotically almost surely. Notably, all firms with a cost price pair (c^*, p^*) are ex ante identical as in our base model. As such, the same argument as in the proof of Theorem 1 guarantees that there exists an equilibrium bid, which in the limit converges to $(p^* - c^*)\left(1 - G(\bar{r} + p^*|\bar{\theta})\right)$. Since there are at least 2 bidders almost surely, the platform's profit equals this bid. This is equal to the upper bound on the platform's profits, implying that the proposed algorithm is asymptotically optimal. Naturally, the match score of the firm that wins the auction approaches $\bar{\theta}$ as $n \to \infty$. □

D Appendix: Proofs for Section 4

Proof of Proposition 5. The proof of this proposition proceeds in three steps:

1. Fix the number of firms n, and suppose the platform after assigning firms to slots reveals the match score of the firm in each slot without the consumer having to inspect a single slot. At this point the consumer can replicate any platform algorithm. Since the consumer can choose the algorithm and her search behavior, the optimal search behavior for the optimal algorithm necessarily achieves the first-best expected consumer surplus. In fact, due to the independence of draws, there is no learning and the optimal consumer search behavior is to calculate reservation values and follow Weitzman's rule.

2. Suppose the platform chooses a perfect ranking but does not reveal match scores. In

this case, the consumer can calculate the expected match score of the firm with the highest match score (the expected maximum of n independent draws from $F(\cdot)$), the second highest and so on. Assume the consumer to use these expected match scores to calculate reservation values, ignoring information from revealed match values. Suppose further that the consumer follows Weitzman's rule, that is, searches in the order of decreasing reservation values. Clearly, if the platform conveys a perfect ranking to the consumer, then expected consumer surplus is bounded from below by the surplus the consumer can achieve via this naive rule.

3. Note that as $n \to \infty$ the realized k-th highest match score converges to the expected $k-th$ highest match score. As a result, as the number of firms n grows large, the expected consumer surplus when calculating reservation values and applying Weitzman's rule with expected match scores approaches the expected consumer surplus when calculating reservation values and applying Weitzman's rule with the realized match scores. As optimal consumer behavior outperforms the former, but the latter achieves the first-best expected consumer surplus, the expected consumer surplus when the firm chooses a perfect ranking approaches the first-best asymptotically.[16] □

Proof of Proposition 6. If higher mark-ups imply lower prices, the price and cost that maximize $(p - c)\,(1 - G(\bar{r} + p|\theta))$ are given by

$$p^* = \min_{h_p(p)>0} p \in \mathcal{P} \qquad \text{and} \qquad c^* = \min_{h_c(c)>0} c \in C.$$

It follows from Proposition 4 (and its proof) that, as $n \to \infty$, the firm winning the sponsored slot will *(i)* feature the cost-price pair (c^*, p^*) a.a.s. and *(ii)* have a match score that is greater than $\theta^* = \bar{\theta} - \varepsilon$ for any $\varepsilon > 0$ a.a.s.. Consequently, since the consumer cares about the match value and the price but not about production cost, the claim follows. □

[16] More formally, one defines a vector with countably infinitely many elements, the first n of which are the realized match scores of ranked firms in decreasing order and zeros thereafter. Invoking an appropriate metric—once then shows this vector to converge to the analogously defined vector of expected match scores of ranked firms. Based on this convergence, one defines the difference between expected consumer surplus from calculating reservation values and invoking Weitzman's rule with expected and realized match scores, and shows that it decreases as $n \to \infty$. We refrain from this level of formality for the sake of parsimony.

Search Platforms: Big Data and Sponsored Positions

Supplementary Material

Maarten Janssen   Thomas Jungbauer   Marcel Preuss   Cole Williams

(FOR ONLINE PUBLICATION)

Lemma S.1. *Let $h : \Omega \to \mathbb{R}_+$ be continuous and $\int h(\omega)d\mu(\omega) = \ell$. Let $\Lambda \subset \Delta\Omega$ be a nonempty subset of probability distributions for which, $\int \min\{h(\omega), m\}d\lambda(\omega) \to \int h(\omega)\lambda(\omega)$ as $m \to \infty$ uniformly for all $\lambda \in \Lambda$. For every neighborhood $\delta > 0$ there exists an $\epsilon > 0$ such that, if $\lambda \in \Lambda$ and $|\lambda(K) - \mu(K)| < \epsilon$ for all measurable K, then $\int h(\omega)\lambda(\omega) \in (\ell - \delta, \ell + \delta)$.*

Proof. To the contrary, suppose that the conclusion does not hold, so that there is a sequence of probability measures $\{\lambda_k\}_{k\in\mathbb{N}}$ satisfying $|\lambda_k(K) - \mu(K)| < 1/k$ for all measurable K and $\int h(\omega)\lambda_k(\omega) \notin (\ell - \delta, \ell + \delta)$ for all $k \in \mathbb{N}$.

Let $0 < \epsilon < \delta$ and define $h_m(\omega) \equiv \min\{h(\omega), m\}$. Let m^* be a value such that, if $m \geq m^*$ then $|\int h(\omega)d\lambda(\omega) - \int h_m(\omega)d\lambda(\omega)| < \epsilon/3$ for all $\lambda \in \Lambda$. Because h_m is bounded and continuous, by weak convergence of $\lambda_k \to \mu$, for every m there is a $k(m)$ such that, if $k \geq k(m)$, then $\left|\int h_m(\omega)d\mu(\omega) - \int h_m(\omega)d\lambda_k(\omega)\right| < \epsilon/3$. By the Monotone Convergence Theorem, there is an m^{**} large enough so that $|\int h(\omega)d\mu(\omega) - \int h_m(\omega)d\mu(\omega)| < \epsilon/3$ whenever $m > m^{**}$. Thus, for $m > \max\{m^*, m^{**}\}$ and $k > k(m)$, by the triangle inequality

$$\left|\int h(\omega)d\mu(\omega) - \int h(\omega)d\lambda_k(\omega)\right|$$

$$\leq \left|\int h(\omega)d\mu(\omega) - \int h_m(\omega)d\mu(\omega)\right| + \left|\int h_m(\omega)d\mu(\omega) - \int h_m(\omega)d\lambda_k(\omega)\right| \quad \text{(S.1)}$$

$$+ \left|\int h_m(\omega)d\lambda_k(\omega) - \int h(\omega)d\lambda_k(\omega)\right| \leq \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon. \quad \text{(S.2)}$$

But this implies a contradiction. □

S.1 Lemma ?? ($\Delta > 0$)

Lemma ?? considers the problem facing an m-restricted consumer at the "limiting distribution" where there is a single sponsored position with a match value distribution $G(v_i|\bar{\theta})$ and $m-1$ organic positions with match value distributions $G(v_i)$. As noted before, the optimal strategy is to follow Weitzman's rule which involves first inspecting the sponsored firm and buying the product whenever $v_S > \bar{r}$. As m grows large, the expected payoff from following this strategy approaches

$$\int_{\underline{v}}^{\bar{v}} \max\{v_S, \bar{r}\} dG(v_S|\bar{\theta}) - s. \tag{S.3}$$

Recall that $\Sigma^m_\delta \subset \Sigma^m$ denotes the subset of strategies for which, conditional on $v_S > \bar{r}$, the limiting distribution places a probability greater than δ that the consumer does not buy from the sponsored firm after the first search. Notice that Σ^m_δ contains two types of strategies, distinguished by whether or not the consumer begins search at the sponsored firm. First, the consumer can begin search at an organic firm; in which case, it is optimal to follow Weitzman's rule thereafter. As m grows large, the expected payoff from this strategy converges to

$$\int \int \left[\max\{v_i, \min\{v_S, r_{\bar{\theta}}\}, \bar{r}\}\right] dG(v_i) dG(v_S|\bar{\theta}) - (1 + G(r_{\bar{\theta}}))s. \tag{S.4}$$

Let $\hat{\Delta} > 0$ denote the difference between (S.3) and (S.4). Second, the consumer can begin search at the sponsored firm. Of these strategies, it is best to follow Weitzman's rule everywhere except when $v_S \in \mathcal{V}$ for some $\mathcal{V} \subset [\bar{r}, \bar{v}]$ satisfying $\mathbb{P}(\mathcal{V}|[\bar{r}, \bar{v}]) = \delta$, where the consumer visits an organic firm and then takes whichever of the two inspected goods delivers a larger value. Computing the expected payoff from using this strategy and subtracting it from the expected from following Weitzman's rule yields a difference of

$$\int \int \max\{0, v_S - v_i\} \mathbf{1}_{\mathcal{V}}(v_S) dG(v_i) dG(v_S|\bar{\theta}) + \delta \cdot s. \tag{S.5}$$

Notice that (S.5) does not depend on m; hence, of these strategies, the optimal $\mathcal{V}$ is the same for all m-restricted consumers. In fact, because $\int \max\{0, v_S - v_i\} dG(v_i)$ is increasing in v_S, the $\mathcal{V}$ that minimizes (S.5) is $[\bar{r}, v^\delta]$ where v^δ is the value that satisfies $\mathbb{P}([\bar{r}, v_\delta]|[\bar{r}, \bar{v}]) = \delta$. Let $\hat{\Delta}'$ be (S.5) evaluated at $\mathcal{V} = [\bar{r}, v^\delta]$.

Recalling the definition $\Delta_m \equiv \lim_{n\to+\infty} \left[\max_{\sigma\in\Sigma^m_\delta} u_{m,n}(\sigma) - \max_{\sigma\in\Sigma^m} u_{m,n}(\sigma)\right]$, it follows from our arguments above that $\lim_{m\to+\infty} \Delta_m = \min\{\hat{\Delta}, \hat{\Delta}'\} > 0$ and thus $\Delta \equiv \inf_{m\in\mathbb{N}} \Delta_m > 0$.

S.2 Example 1

In this subsection, we discuss an example illustrating how uniform obfuscation can fail to be optimal when there are a small number of firms, because of non-monotonicities in the inference consumers draw from observing their match value at the sponsored slot.

Suppose there are three firms $i = 1, 2, 3$. The consumer's match value is either low ℓ, medium m, or high h and a good is only worth purchasing if it provides at least a medium value. A firm's match score is L when the value is low and H when the value is either medium or high, i.e., the platform can distinguish firms with low match scores from other firms, but cannot distinguish firms with medium and high match scores. Let p_H and p_L denote the marginal probability that a firm's score is high and low, respectively.[17]

Suppose the platform employs the following algorithm. The firm with the highest bid is placed in sponsored positions, ties are broken in favor of the firm with the highest match score, further ties are broken with equal probability. For the two nonsponsored firms, if only one of them has a high signal it is placed in the second position with probability $\alpha \geq \frac{1}{2}$, otherwise they are arranged in the organic positions with equal probability.

Given the algorithm, the consumer's optimal search proceeds in the following manner. If the sponsored firm's value is high h, then the consumer buys it immediately since there is no advantage from continuing. If instead the sponsored firm's value is low ℓ, then given the algorithm, the consumer learns that all remaining firms must likewise have low match values and so the consumer might as well exit the market. If, however, the consumer observes m in the sponsored slot, then it might still be prudent to continue searching as some remaining firm might deliver a higher match value. To describe the consumer's learning over the course of search, let subscripts denote the index of the position so that the list of possible events are $\{(H_1, H_2, H_3), (H_1, L_2, H_3), (H_1, H_2, L_3), (H_1, L_2, L_3)(L_1, L_2, L_3)\}$ which occur with corresponding probabilities $\{p_H^3, 3(1-\alpha)p_H^2 p_L, 3\alpha p_H^2 p_L, 3p_H p_L^2, p_L^3\}$. The probability

[17]This example departs from the assumptions of our model in that the distribution of match values conditional on the match scores do not share the same support. This is insignificant to the particular example since we could modify the distributions to $\mathbb{P}(\{\ell\}|L) = \mathbb{P}(\{m, h\}|H) = 1 - \varepsilon$ so that the conclusion continues to hold for $\varepsilon > 0$ sufficiently small.

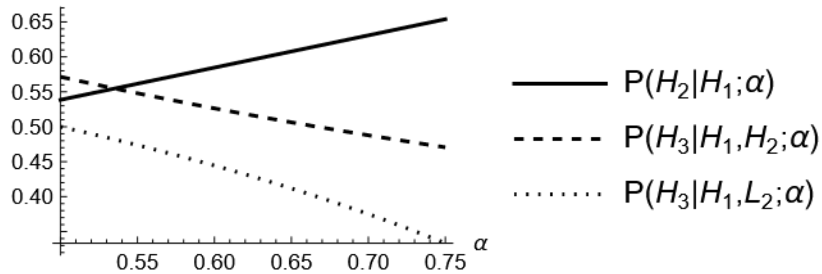


Figure 1: The figure plots the conditional probabilities as a function of α given that match values each occur with equal probability.

that slot two has a high score given that the first does is

$$\mathbb{P}(H_2|H_1;\alpha) = \frac{p_H^3 + 3\alpha p_H^2 p_L}{1 - p_L^3}. \tag{S.6}$$

The probability that slot three has a high score given that the first two also do is

$$\mathbb{P}(H_3|H_1,H_2;\alpha) = \frac{p_H^3}{p_H^3 + 3\alpha p_H^2 p_L}. \tag{S.7}$$

The probability that slot three has a high score given that the first does and the second has a low score is

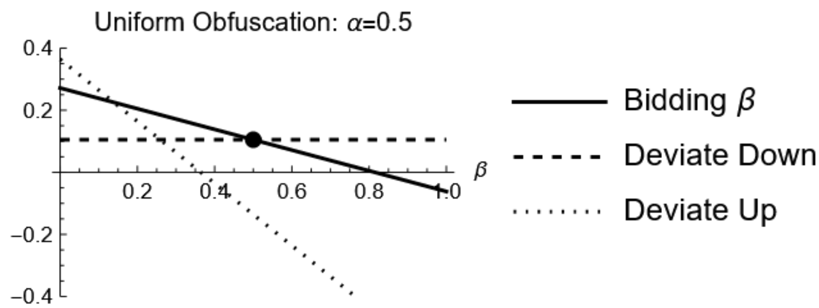
$$\mathbb{P}(H_3|H_1,L_2;\alpha) = \frac{3(1-\alpha)p_H^2 p_L}{3(1-\alpha)p_H^2 p_L + 3p_H p_L^2}. \tag{S.8}$$

Inspecting the above expressions, we find that (S.7)>(S.8) which naturally implies that the consumer is more optimistic about the third slot upon observing a medium match value in the first two slots than if she were to observe a medium in the first and a low in the second. Also, (S.6)>(S.8) holds true, implying the consumer is more optimistic about the second slot after observing a medium in the first than she is about the third slot upon observing a medium and low value in the first two. The comparison between (S.6) and (S.7) depends on α. As Figure 1 illustrates, increasing α makes proceeding to the second firm more attractive, but continuing to the third less so.

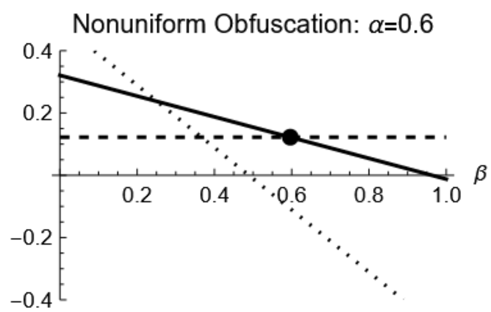
Suppose the parameters are such that, under uniform obfuscation (i.e. when $\alpha = 1/2$), the consumer continues searching when observing m in the first firm and m in the second firm, but halts otherwise. To be concrete, assume that the parameters satisfy $\frac{1}{2}\mathbb{P}(H_2|H_1;\alpha = \frac{1}{2})(h-m) = s$ which leads the consumer to follow the desired search pattern. We compare uniform obfuscation against a nonuniformly obfuscating algorithm with $\alpha = \alpha^* > \frac{1}{2}$ whereby α^* is large enough to ensure that a consumer will inspect the second slot if the first provides a medium match value, but will search no further. For example, in Figure 1, given our assumptions on parameters, setting $\alpha = 0.6$ guarantees that $\mathbb{P}(H_3|H_1, H_2; \alpha^*) < \mathbb{P}(H_2|H_1; \frac{1}{2})$ and thus it is never optimal for the consumer to inspect the third firm. Notice that by providing some information in the organic slots, the nonuniformly obfuscating algorithm makes inspecting the second firm more desirable, but increases a sponsored firm's return demand as consumers will not inspect the third firm.

For each algorithm, consider a symmetric equilibrium in which each firm bids β. Figure 2 plots the tentative expected equilibrium and deviation profits under the two proposed algorithms for different values of the bid β. Naturally, the expected profit from playing the tentative equilibrium strategy is decreasing in the bid with a slope of $-\frac{1}{3}$ as each of the three firms win with equal probability. Placing a higher bid, no matter how high, ensures that a firm wins the auction, but also makes winning uninformative as it does not provide information about the firm's match score. The slope of the expected profit given an upward deviation is thus -1. Thus, this deviation risks winning in case the firm has a medium match score and consumers continue searching after visiting the sponsored slot. The deviation is optimal for low values of the candidate equilibrium bid, but not for higher values. Offering a bid less than β ensures that a firm does not win the sponsored slot; hence, the deviation profit is simply the expected profit in an organic slot and does not depend on the bid.

We can use Figure 2 to compare the profitability of the two algorithms by comparing the range of bids firms are willing to place. The figure shows that, in principle, there can be a continuum of equilibrium bids. However, we only need to compare the equilibria with the highest bids as the platform can secure a profit equal to this bid by setting a reserve price equal to the highest possible equilibrium bid. As illustrated by Figure 2, using a uniform obfuscation algorithm, the platform can achieve a profit of exactly $\frac{1}{2}$, which is identified by finding the highest value for the bid at which no firm wishes to deviate from placing that bid. On the other hand, using the nonuniformly obfuscating algorithm, the platform can secure a profit of approximately 0.6. Thus, due to the learning from match values at the sponsored slots, the platform is better off choosing the nonuniformly obfuscating algorithm.



(a)



(b)

Figure 2: The figures plot an individual firm's expected profit in a tentative symmetric equilibrium in which all firms bid β under the respective algorithms from likewise bidding β, deviating to a lower bid, and deviating to a higher bid (given that match values each occur with equal probability).

S.3 Example 2

In this subsection, we discuss an example illustrating how the platform may deviate from assigning the firm with the highest match score to the sponsored slot when there are a small number of firms. To this end, consider a similar setup as before in Example 1 where there are three distinct match values l, m, h and where the platform cannot distinguish medium and high values, but perfectly recognizes a low value firm. Thus, when the match value is l, the match score is L and when the match values is either m or h the platform has a match score of H. Now there are four firms, however. Moreover, suppose the consumer would choose any match value, including the lowest l, over her outside option.

Consider two different algorithms, one where the platform always puts the firm with the highest match score in the sponsored position and uniformly obfuscates the organic positions and another, second, algorithm where the platform almost always follows the same algorithm apart from the case where two match scores are L and two are H. In that case the platform puts a firm with the lowest match score in the sponsored position and uniformly obfuscates the organic positions.

We show the conditions under which the second algorithm gives the winning firm a higher probability of selling. In particular, consider that with the first algorithm the consumer buys immediately from the sponsored position whenever it contains an h or l value, but he continues to search if he sees an m value. On the other hand, with the second algorithm the consumer buys immediately from the sponsored position whatever its match value. As the second algorithm makes sure that the consumer will always buy from the sponsored position and never buys from an organic position, it is clear that platform gets its highest possible profit as firms bid maximally to get into the sponsored position.

We first consider under what conditions after observing an m in the sponsored slot the consumer continues to search under the first algorithm, but not under the second, and we start the analysis with the second algorithm. We use backard induction to determine the pay-off for the consumer *if* he decides to continue to search after observing an m. So, consider that after observing an m in the first round, the consumer observes either ll, ml (or lm) or mm in the subsequent two rounds (and we thus focus on whether the consumer wants to inspect the last object or not). If, in the meantime a consumer has found an h object, he will of course buy immediately).

(i) After observing mll in the first three search rounds, the consumer obviously stops searching, as he must update his beliefs in such a way that the last object is also an l. His pay-off in this case is $m - 3s$.

(ii) After observing mlm or mml in the first three search rounds, we postulate the consumer continues searching as under the second algorithm the only pattern that is consistent with this is for there to be an H object in the last round. Thus, the consumer's pay-off, then would be $\frac{1}{2}(m+h) - 4s > m - 3s$. Thus, we implicitly assume that $h - m > 2s$.

(iii) Finally, after observing mmm in the first three search rounds, we postulate the consumer continues searching. Updatiing his beliefs, if the consumer continues to search he encounters an L on the last search with probability $\frac{\frac{1}{3}\cdot 4p_H^3 p_L}{\frac{1}{3}\cdot 4p_H^3 p_L + p_H^4} = \frac{4p_L}{4p_L+3p_H}$ and an H with probability $\frac{3p_H}{4p_L+3p_H}$. Thus, the consumer's pay-off would then be $\frac{8p_L+3p_H}{8p_L+6p_H}m + \frac{3p_H}{8p_L+6p_H}h - 4s$ (and this is better than stop searching with a payoff of $m - 3s$, if $\frac{3p_H}{8p_L+6p_H}(h-m) > s$). Thus, we implicitly assume that

$$\frac{h-m}{s} > 2 + \frac{8p_L}{3p_H}. \tag{S.9}$$

Note that this condition also implies that $h - m > 2s$.

Let us then go back one search round and consider that the consumer after observing an m in the first round, he observes either l or m in the next round.

(i) After observing ml in the first two search rounds, we postulate the consumer stops searching, as he must update his beliefs in such a way that the last two objects are either ll or HH. Given the second algorithm, the total ex ante probability that the consumer observes ml in the first two search rounds is $\frac{1}{3}\cdot 4p_H^3 p_L + 4p_L^3 p_H$ and thus continuing to search yields an expected pay-off of

$$\frac{4p_L^3 p_H}{\frac{1}{3}\cdot 4p_H^3 p_L + 4p_L^3 p_H}(m-3s) + \frac{\frac{1}{3}\cdot 4p_H^3 p_L}{\frac{1}{3}\cdot 4p_H^3 p_L + 4p_L^3 p_H}\left(\frac{1}{2}(h-3s) + \frac{1}{2}\left(\frac{1}{2}(m+h) - 4s\right)\right).$$

This is smaller than $m - 2s$ (the pay-off if he stops searching) if, and only if, $\frac{1}{2}(h-2s) + \frac{1}{2}\left(\frac{1}{2}(m+h) - 3s\right) < m - 2s + \frac{\frac{1}{3}\cdot 4p_H^3 p_L + 4p_L^3 p_H}{\frac{1}{3}\cdot 4p_H^3 p_L}s$, or $\frac{3}{4}(h-m) < \left(\frac{3}{2} + \frac{3p_L^2}{p_H^2}\right)s$, which is true if

$$\frac{h-m}{s} < 2 + \frac{4p_L^2}{p_H^2}. \tag{S.10}$$

(ii) After observing mm in the first two search rounds, we postulate the consumer continues to search all the way until he found an h object, as he must update his beliefs in such a way that the last two objects are either HL, LH or HH (which happens with a total ex ante probability of $\frac{2}{3} \cdot 4p_H^3 p_L + p_H^4$) and thus he gets an additional pay-off of continuing to search of

$$\begin{aligned}
& \frac{p_H}{p_H + \frac{2}{3} \cdot 4p_L} \left(\frac{1}{2} (h - m - s) + \frac{1}{2} \left(\frac{1}{2}(h - m) - 2s \right) \right) + \frac{\frac{2}{3} \cdot 4p_L}{p_H + \frac{2}{3} \cdot 4p_L} \left(\frac{1}{2}(h - m - \frac{1}{2}s) - \frac{3}{4} \cdot 2s \right) \\
= \; & \frac{3p_H}{3p_H + 8p_L} \left(\frac{3}{4}(h - m) - 1\frac{1}{2}s \right) + \frac{8p_L}{3p_H + 8p_L} \left(\frac{1}{2}(h - m) - \frac{7}{4}s \right) \\
= \; & \frac{\frac{9}{4}p_H + 4p_L}{3p_H + 8p_L}(h - m) - \frac{\frac{9}{2}p_H + 14p_L}{3p_H + 8p_L}s,
\end{aligned}$$

which should be larger than 0 for him to prefer to continue searching, which is the case if

$$\frac{h - m}{s} > \frac{\frac{9}{2}p_H + 14p_L}{\frac{9}{4}p_H + 4p_L}. \tag{S.11}$$

So, now we can give a condition under which the consumer stops searching after observing an m in the first slot as his overall additional pay-off (incorporating learning and optimal search) of continuing to search after observing an m in the first round instead of stopping immediately is (as the overall probability of this event happening is $p_H^4 + 4p_H^3 p_L + 4p_L^3 p_H$):

$$\begin{aligned}
& -\frac{4p_L^3 p_H + \frac{1}{3} \cdot 4p_H^3 p_L}{p_H^4 + 4p_H^3 p_L + 4p_L^3 p_H} s + \\
& \frac{p_H^4 + \frac{2}{3} \cdot 4p_H^3 p_L}{p_H^4 + 4p_H^3 p_L + 4p_L^3 p_H} \left\{ \frac{1}{2}(h - m) + \frac{1}{2} \left(\frac{\frac{9}{4}p_H + 4p_L}{3p_H + 8p_L}(h - m) - \frac{\frac{9}{2}p_H + 14p_L}{3p_H + 8p_L}s \right) - s \right\} < 0,
\end{aligned}$$

which after combining terms yields

$$p_H^2(3p_H + 8p_L)\frac{h - m}{6} \left(1 + \frac{\frac{9}{4}p_H + 4p_L}{3p_H + 8p_L} \right) < \left(\frac{1}{3}p_H^2(3p_H + 8p_L) \left(1 + \frac{\frac{9}{4}p_H + 7p_L}{3p_H + 8p_L} \right) + 4p_L \left(p_L^2 + \frac{1}{3}p_H^2 \right) \right) s,$$

or

$$3p_H^2 \frac{h - m}{24} (7p_H + 16p_L) < \left(\frac{1}{3}p_H^2 \left(\frac{21}{4}p_H + 19p_L \right) + 4p_L^3 \right) s,$$

or

$$\frac{h - m}{s} < \frac{42p_H^3 + 152p_H^2 p_L + 96p_L^3}{21p_H^3 + 48p_H^2 p_L} \tag{S.12}$$

Let us now do the same exercise for the first algorithm and use backward induction to determine the pay-off for the consumer *if* he decides to continue to search after observing an m. So, consider that after observing an m in the first round, the consumer observes either ll, ml (or lm) or mm in the subsequent two rounds.

(i) After observing ll in search rounds two and three, the consumer encounters an L with probability $\frac{4p_L^3 p_H}{4p_L^3 p_H + 2p_H^2 p_L^2} = \frac{2p_L}{2p_L + p_H}$ on the last search and an H with probability $\frac{p_H}{2p_L + p_H}$.

(ii) After observing lm (the same analysis applies to ml) in search rounds two and three, the total ex ante probability of $HHLH$ and $HHLL$ is $\frac{1}{3} \cdot 4p_H^3 p_L + \frac{1}{3} \cdot 6p_H^2 p_L^2$, where for example $6p_H^2 p_L^2$ is the ex ante probability that there are two H and two L and in that case (under the first algorithm) the sponsored slot is an H and the chance that the first organic slot contains an L is 1/3. So the conditional probability of the last one being H, resp. L, is $\frac{4p_H^3 p_L}{4p_H^3 p_L + 6p_H^2 p_L^2} = \frac{2p_H}{2p_H + 3p_L}$ and $\frac{3p_L}{2p_H + 3p_L}$.

(iii) Finally, after observing mm in search rounds two and three, the total ex ante probability of $HHHH$ and $HHHL$ is $p_H^4 + \frac{1}{3} \cdot 4p_H^3 p_L$ so the conditional probability of the last one being H, resp. L, is $\frac{p_H}{p_H + \frac{4}{3}p_L}$ and $\frac{\frac{4}{3}p_L}{p_H + \frac{4}{3}p_L}$.

If the consumer continues to search in the first case, he certainly continues to search in all other cases and this is the case if $\frac{p_H}{p_H + 2p_L}\frac{m+h}{2} + \frac{2p_L}{p_H + 2p_L} m - s > m$, or $\frac{p_H}{p_H + 2p_L}\frac{h-m}{2} > s$ or

$$\frac{h-m}{s} > 2\left(1 + \frac{2p_L}{p_H}\right). \tag{S.13}$$

It is clear that (S.13) implies (S.9) and (S.11).

What we will do in the subsequent analysis is the following. We show that if (S.13) holds, then the consumer wants to continue searching in all previous search rounds until he has found an h. The easiest way to do so is to show that if (S.13) holds the consumer prefers to search in round t even if in subsequent round $t + 1$ he stops searching. As the pay-off of continuing to search in round t is higher than that (as the consumer actually continues to search as this yields a higher pay-off), he certainly wants to continue searching if that higher continuation pay-off is taking into account.

With this in mind, let us then go back one period and consider that the consumer after observing an m in the first round either observes l or m in the next round.

(i) After observing l in search round two, the consumer believes that the last two objects are either LL, HL or HH and so overall, he believes the objects are either $HLHH, HLHL, HLLH$ or $HLLL$ and the ex annte total probability that one of these events happens is $\frac{1}{3} \cdot 4p_H^3 p_L + \frac{2}{3} \cdot 6p_H^2 p_L^2 + 4p_H p_L^3$.(For example, the ex ante probability that three products are H and one is L is $4p_H^3 p_L$ and out of these cases under the first algorithm the probability that the first objects ranked is an H and the next one is an L is $\frac{1}{3}$.) Thus, the conditional probability that the third object searched is an H equals $\frac{\frac{1}{3}\cdot 4p_H^3 p_L + \frac{1}{3}\cdot 6p_H^2 p_L^2}{\frac{1}{3}\cdot 4p_H^3 p_L + \frac{2}{3}\cdot 6p_H^2 p_L^2 + 4p_H p_L^3}$ and therefore the consumer prefers searching in round 3 even if he does not continue searching in round 4 if

$$\frac{\frac{1}{3} \cdot 4p_H^3 p_L + \frac{1}{3} \cdot 6p_H^2 p_L^2}{\frac{1}{3} \cdot 4p_H^3 p_L + \frac{2}{3} \cdot 6p_H^2 p_L^2 + 4p_H p_L^3} \frac{h-m}{2} > s.$$

This condition can be rewritten as $\frac{h-m}{s} > 2\frac{2p_H^2 + 6p_H p_L + 6p_L^2}{2p_H^2 + 3p_H p_L} = 2\left(1 + \frac{3p_H p_L + 6p_L^2}{2p_H^2 + 3p_H p_L}\right)$. The RHS of this inequality is smaller than the RHS of (S.13) if $\frac{3p_H + 6p_L}{2p_H + 3p_L} < 2$, which is clearly the case.

(ii) After observing m in search round two, the ex ante total probability of the first two objects being H equals $p_H^4 + \frac{2}{3} \cdot 4p_H^3 p_L + \frac{1}{3} \cdot 6p_H^2 p_L^2$. Thus, the conditional probability that the third object searched is an H equals $\frac{p_H^4 + \frac{1}{3}\cdot 4p_H^3 p_L}{p_H^4 + \frac{2}{3}\cdot 4p_H^3 p_L + \frac{1}{3}\cdot 6p_H^2 p_L^2}$ and therefore the consumer prefers searching in round 3 even if he does not continue searching in round 4 if

$$\frac{p_H^4 + \frac{1}{3} \cdot 4p_H^3 p_L}{p_H^4 + \frac{2}{3} \cdot 4p_H^3 p_L + \frac{1}{3} \cdot 6p_H^2 p_L^2} \frac{h-m}{2} > s,$$

which can be written as $\frac{h-m}{s} > 2\frac{3p_H^2 + 8p_H p_L + 6p_L^2}{3p_H^2 + 4p_H p_L} = 2\left(1 + \frac{4p_H p_L + 6p_L^2}{3p_H^2 + 4p_H p_L}\right)$. The RHS of this inequality is smaller than the RHS of (S.13) if $\frac{4p_H + 6p_L}{3p_H + 4p_L} < 2$, which again is clearly the case.

So, now we go to the first search round where the consumer observes an m. The ex ante total probability of this state equals $p_H^4 + 4p_H^3 p_L + 6p_H^2 p_L^2 + 4p_H p_L^3$. Thus, the conditional probability that the second object searched is an H equals $\frac{p_H^4 + \frac{2}{3}\cdot 4p_H^3 p_L + \frac{1}{3}\cdot 6p_H^2 p_L^2}{p_H^4 + 4p_H^3 p_L + 6p_H^2 p_L^2 + 4p_H p_L^3}$ and therefore the consumer prefers searching in round 2 even if he does not continue searching in round 3 if

$$\frac{p_H^4 + \frac{2}{3} \cdot 4p_H^3 p_L + \frac{1}{3} \cdot 6p_H^2 p_L^2}{p_H^4 + 4p_H^3 p_L + 6p_H^2 p_L^2 + 4p_H p_L^3} \frac{h-m}{2} > s,$$

which can be written as $\frac{h-m}{s} > 2\frac{p_H^4 + 4p_H^3 p_L + 6p_H^2 p_L^2 + 4p_H p_L^3}{p_H^4 + \frac{2}{3}\cdot 4p_H^3 p_L + \frac{1}{3}\cdot 6p_H^2 p_L^2} = 2\left(1 + \frac{\frac{1}{3}\cdot 4p_H^3 p_L + \frac{2}{3}\cdot 6p_H^2 p_L^2 + 4p_H p_L^3}{p_H^4 + \frac{2}{3}\cdot 4p_H^3 p_L + \frac{1}{3}\cdot 6p_H^2 p_L^2}\right)$. The RHS of this inequality is smaller than the RHS of (S.13) if $\frac{\frac{1}{3}\cdot 4p_H^3 + \frac{2}{3}\cdot 6p_H^2 p_L + 4p_H p_L^2}{p_H^3 + \frac{2}{3}\cdot 4p_H^2 p_L + \frac{1}{3}\cdot 6p_H p_L^2} < 2$, which again is clearly the case.

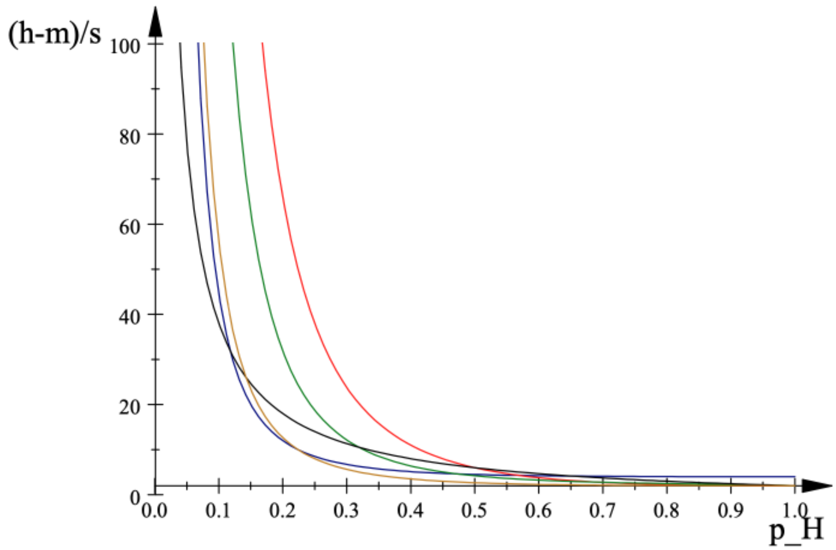


Figure 3

Thus, combining (S.10), (S.12) and (S.13), a sufficient condition for it to be possible that after observing m the consumer continues searching after the first algorithm, but not under the second is that

$$2\left(1+\frac{2p_L}{p_H}\right) < \frac{h-m}{s} < \min\left\{\frac{42p_H^3+152p_H^2p_L+96p_L^3}{21p_H^3+48p_H^2p_L}, 2+\frac{4p_L^2}{p_H^2}\right\}. \tag{S.14}$$

As the LHS and the RHS are function of p_H only (as $p_L = 1 - p_H$), it is possible to choose values of $h - m, s$ such that both inequalities hold if there exist values of p_H such that the RHS is larger than the left hand side. Figure 3 draws $\frac{h-m}{s}$ on the vertical axis and p_H on the horizontal axis. The term of the LHS is represented by the black curve and the two terms of the RHS are represented by the green and red curve, respectively. It is clear that for (roughly) $p_H < 0.27$ the black curve is below the green curve and the green curve is the constraining factor on the RHS, reflecting the first term on the RHS. Thus, for the area in between the black and the green curve inequality (S.14) holds.

Finally, we consider the conditions under which it is optimal to stop searching under the second algorithm after observing an l on the first search. (It is clear it is optimal to stop searching after observing an l under the first algorithm). After observing an l under the second algorithm, the consumer knows there are either two H products or all are L. It is clear that after observing three l the consumer stops searching. Also, after observing two L (which happens with a total ex ante probability of $\frac{1}{3} \cdot 6p_H^2 p_L^2 + p_L^4$) the consumer knows the continuation is either LL or HH and then the consumer stops searching if

$$\frac{2p_H^2}{2p_H^2 + p_L^2}\left((\frac{m+h}{2} - s) + \frac{1}{2}(\frac{h-m}{2} - s)\right) + \frac{p_L^2}{2p_H^2 + p_L^2}(l - s) < l,$$

or if $l \approx m$

$$\frac{(h-m)}{s} < 2 + \frac{2p_L^2}{3p_H^2}, \tag{S.15}$$

which is depicted by the beige curve in the figure.

Finally, after observing lm the consumer knows that there is still one L and one H option and it is then optimal to continue to search if the H option on the second search turned out to be an m if $\frac{1}{2}\left(\frac{1}{2}(h-m) - s\right) + \frac{1}{2}\left(\frac{1}{2}(h-m) - 2s\right) > 0$, which is the case if $\frac{h-m}{s} > 3$, which is clearly implied by the previous conditions in the figure. (The pay-off formula follows from the fact that with probability $\frac{1}{2}$ the next option is an H and then the consumer certainly stops searching afterwards, while with the remaining probability $\frac{1}{2}$ the next option is an L and then the consumer certainly continues searching).

Thus, after observing an L in the first position, the pay-off of continuing to search is given by

$$\frac{2p_H^2 + p_L^2}{6p_H^2 + p_L^2}(l - s) + \frac{4p_H^2}{6p_H^2 + p_L^2}\left(\frac{1}{3}(l - s) + \frac{1}{3}(h - s) + \frac{1}{3}(m + \frac{1}{2}(h - m) - 2\frac{1}{2}s)\right).$$

For $l \approx m$ this is smaller than l if (approx.)

$$\frac{4p_H^2}{6p_H^2 + p_L^2}\left(\frac{1}{3}(m - s) + \frac{1}{3}(h - s) + \frac{1}{3}(\frac{1}{2}(m + h) - 2\frac{1}{2}s)\right) < \frac{4p_H^2}{6p_H^2 + p_L^2}m + \frac{2p_H^2 + p_L^2}{6p_H^2 + p_L^2}s,$$

or

$$\frac{4p_H^2}{6p_H^2 + p_L^2}\left(-\frac{3}{2}s + \frac{1}{2}(h - m)\right) < \frac{2p_H^2 + p_L^2}{6p_H^2 + p_L^2}s,$$

or

$$\frac{h-m}{s} < 3 + \frac{p_L^2 + 2p_H^2}{2p_H^2} = 4 + \frac{p_L^2}{2p_H^2}, \tag{S.16}$$

which is depicted by the blue curve in the graph.

Thus, we can conclude that there is a parameter region where the stipulated search behavior under both algorithms is optimal at every stage, which is given by the black curve as the lower bound and the blue curve as the upper bound. In this region, the consumer always stops searching at the sponsored slot under the second algorithm but continues to search after observing an m in the sponsored slot under the first algorithm. It is clear from the analysis that we have provided sufficient conditions for this to be the case that are by no means necessary conditions. The total area of parameter values where the consumer always stops searching at the sponsored slot under the second algorithm but continues to search after observing an m in the sponsored slot under the first algorithm is larger than the area between the black and blue curves depicted in the figure.

S.4 Common Support Assumption

This section presents a counterexample which demonstrates that the conclusion of Theorem 1 no longer holds if we remove the assumption that the conditional distributions $G(v_i|\theta_i)$ share a common support for each score $\theta_i \in [\underline{\theta}, \bar{\theta}]$.

Consider the example in §S.2, except now assume that the consumer is willing to purchase for any match value, though of course still prefers a larger value. With this modification, in a tentative equilibrium with uniform obfuscation, observing that the sponsored firm has a low value still discourages the consumer from inspecting any organic firms since she believes that they all must likewise have a low match value. However, unlike in the earlier example, the discouraged consumer now buys the low value good from the sponsored firm. If we suppose that a consumer is willing to continue searching organic firms when the sponsored firm's value is medium, then letting the number of firms grow large, the expected demand for a sponsored firm given that it has played the equilibrium strategy is $\frac{1}{2}$. That is, in the limit, when playing the tentative equilibrium strategy, the sponsored firm almost certainly has a high signal and thus only makes a sale if it provides a high value. However, if we consider the demand for a firm that deviates to a higher bid and secures the sponsored position, then its demand is $\frac{1}{2}p_H + p_L$ since it makes a sale if it offers either a high or low value. This contradicts Lemma B.3 and also shows that the demand for a lower-valued firm could exceed that of a higher-valued one.